IRE Transactions

n Microwave Theory and Techniques

Volume MTT-8 **MAY, 1960** Number 3

In This Issue

PUBLISHED BY THE

ofessional Group on Microwave Theory and Techniques

IRE TRANSACTIONS®
on Microwave Theory and Techniques

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Microwave Theory and Techniques, at 1 East 79th Street, New York 21, New York. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group, or its members. Price per copy: IRE PGMTT members, $1.95; IRE members, $2.95, nonmembers, $5.85. Annual subscription price: IRE members, $8.50; colleges and public libraries, $12.75; nonmembers, $17.00.

Address all manuscripts to Donald D. King, PGMTT Editor, Electronic Communications, Inc., 1830 York Road, Timonium, Md. Submission of three copies of manuscripts, including figures, will expedite the review.

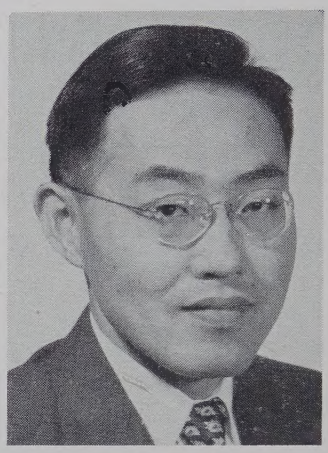
IRE Transactions on Microwave Theory and Techniques

Volume MTT-8 MAY, 1960 Number 3

TABLE OF CONTENTS

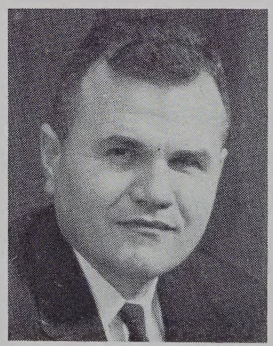
Professional Group on Microwave Theory and Techniques Administrative Committee

A. A. Oliner
Chairman



K. Tomiyasu
Vice Chairman, Meetings Committee Chairman

H. M. Altschuler
Secretary-Treasurer



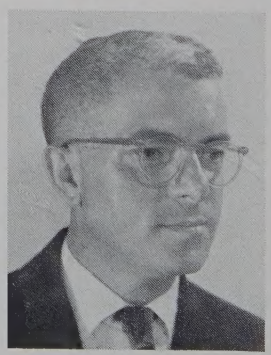
T. N. Anderson
Advertising Editor, Chapter Activities Chairman

R. E. Beam

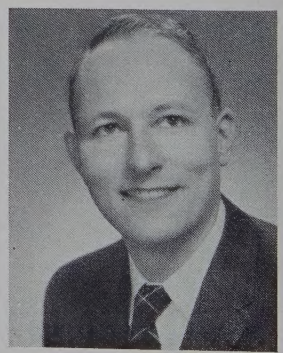
A. C. Beck
Nominations Committee Chairman

A. G. Clavier

S. B. Cohn

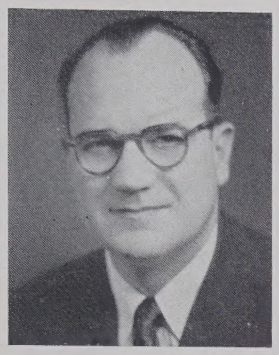


H. F. Engelmann
Ex-Officio

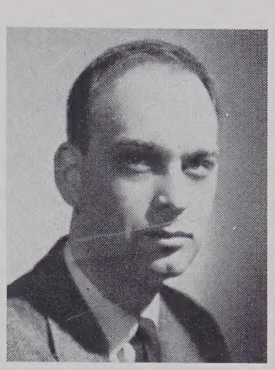


R. C. Hansen

D. D. KING
Transactions Editor

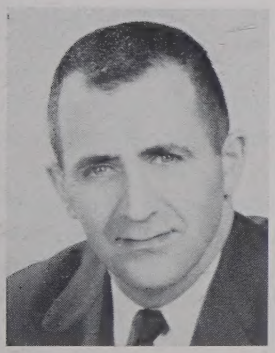


W. W. MUMFORD
Awards Committee Chairman

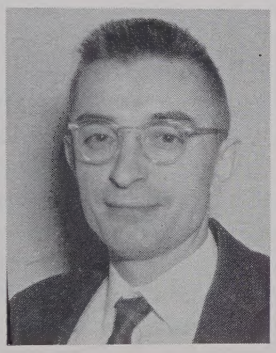


W. L. PRITCHARD
Senior Past Chairman, Anniversary Issue Committee Chairman

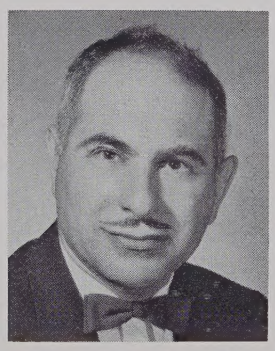
S. W. ROSENTHAL



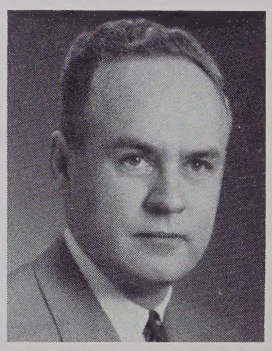
T. S. SAAD
Junior Past Chairman



R. F. SCHWARTZ
Membership Chairman

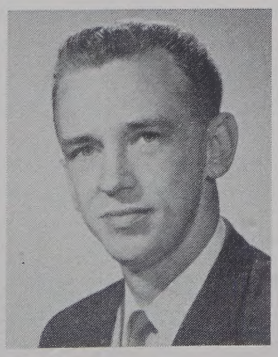


G. SHAPIRO
Newsletter Editor

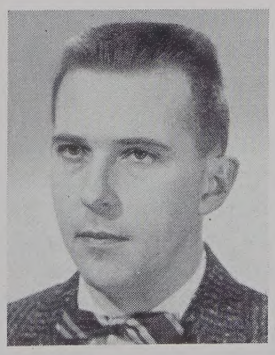


G. SINCLAIR

G. C. SOUTHWORTH
Honorary Life Member



P. D. STRUM



M. C. THOMPSON, JR.

R. D. WENGENROTH

An Analysis of Four-Frequency Nonlinear Reactance Circuits*

DAVID K. ADAMS†

Summary—Several advantages of multiple-frequency nonlinear reactance circuits are described in this paper. In particular, a circuit is considered in which a nonlinear reactance couples four basic frequencies: ω_0, ω_1, ω_2, and ω_3; these are so related that $\omega_2=\omega_0+\omega_1$ and $\omega_3=\omega_0-\omega_1$. Here, ω_0 is taken to be the power source or pump. It is found to be desirable to allow for the possible presence of the pump harmonic, $2\omega_0$, and individual cases are characterized by whether $2\omega_0$ is present or not. The major results are as follows: 1) Unlimited amplification gain is theoretically possible at frequencies higher than the pump, by reflecting negative input resistance at ω_2, but without relying on any effects due to pump harmonics. 2) Unlimited up- or down-conversion gains between ω_1 and ω_2 are theoretically possible in the additional presence of the first pump harmonic, but without reflecting negative input or output resistance. 3) Unlimited amplification gain is theoretically possible at frequencies both lower and higher than the pump fundamental, without reflecting negative input resistance.

Introduction

In recent years, considerable attention has been given to the unusual properties of nonlinear reactance modulators. The basic attribute of these circuits that has been exploited thus far is frequency-conversion with power gain. Since nonlinear reactance conversion gain can be obtained in the presence of a signal (or reference frequency), a local oscillator (or "pump"), and just one of their sidebands, emphasis to date has been on three-frequency circuits. It has been demonstrated, however, that remarkably different conversion characteristics arise, depending upon the particular choice of sideband.[1] In the upper-sideband case, the magnitude of conversion gain is limited to the ratio of the output frequency to the input frequency. In the lower-sideband case, arbitrary conversion gain is possible through regenerative action, which reflects negative resistance into both the input and the output terminals. Such negative resistance also enhances the input signal; this effect is known as parametric amplification.

In view of the striking differences between these two sidebands, the simultaneous presence of both sidebands causes the anticipation of interesting effects. Several of these effects will be investigated in this paper, and the following major results will be demonstrated:

1) Parametric amplification is possible at frequencies higher than the pump fundamental, without relying upon effects caused by pump harmonics.

2) In the additional presence of the first pump harmonic, arbitrary up- or down-conversion gain is possible between the reference frequency and the upper sideband without reflecting negative resistance at either frequency. Furthermore, the reflected input and output resistances can be made insensitive to gain.

With nonlinear reactance elements possessing strong nonlinearities and good high-frequency response, the value of the first result lies primarily in its contribution to the general understanding of nonlinear reactance circuits. However, the second result has considerable potential value in spite of its dependence on the presence of the first pump harmonic. As a consequence of statement 2), up- or down-converters of arbitrary gain can be cascaded. In the special case in which an up-converter of the type mentioned above is followed by a similar down-converter, a nonlinear reactance amplifier can be constructed which does not rely on the reflection of negative resistance at the input terminals. Therefore, unlike conventional parametric amplifiers and negative-input-resistance converters, the full potential of this improved converter circuit can be obtained without the use of circulators.

General Energy Considerations

The basic problem to be considered in the following analysis is that of a lossless, nonlinear reactance in the presence of signals at ω_0, ω_1, and at their upper and lower sidebands: $\omega_2=\omega_0+\omega_1$ and $\omega_3=\omega_0-\omega_1$. In general, the signal at ω_0 will be considered as the power source or pump and will be assumed to have arbitrarily large available power. It is practical to allow also for the possible presence of the pump harmonic, $2\omega_0$, since it will be shown to provide valuable effects in some cases.

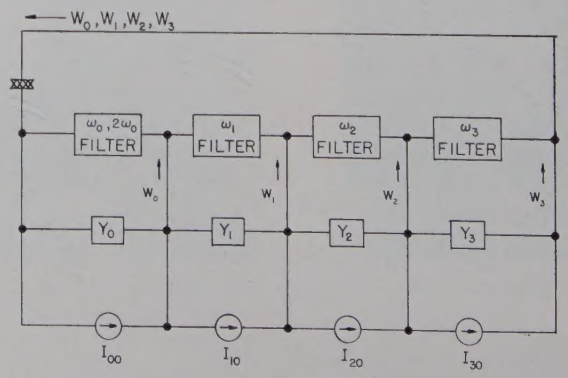


Fig. 1—A general circuit model for the nonlinear reactance problem under discussion.

* Manuscript received by the PGMTT, July 31, 1959; revised manuscript received, December 18, 1959. This work was sponsored by the U. S. Army Signal Res. and Dev. Lab., Ft. Monmouth, N. J., under Contract No. DA-36-039 sc 78283. The conclusions of this paper were originally reported at the PGMTT National Symposium, Harvard University, Cambridge, Mass., June, 1959.

† Electronic Defense Group, University of Michigan Research Institute, Ann Arbor, Mich.

[1] J. M. Manley and H. E. Rowe, "Some general properties of nonlinear elements—Part I. General energy relations," PROC. IRE, vol. 44, pp. 904–913; July, 1956.

A circuit model for this problem is suggested in Fig. 1, in which a nonlinear capacitor has been chosen to represent the nonlinear reactive element, and ideal filters have been assumed: 1) to short circuit all frequencies except those chosen above, and 2) to allow these frequencies to be coupled only by the nonlinear capacitor. While the assumption of ideal filters seems to be very restricting, the external circuits will ultimately be constrained to be resonant with fairly high Q's. Thus, the ideal filters can eventually be removed without seriously affecting circuit operation.

Although small-signal analysis will be the primary medium employed in this investigation, the general properties of nonlinear reactance elements provide a foundation for the results cited previously. It has been shown by Manley and Rowe[1] that the following general statements can be made about lossless, nonlinear reactances in the presence of sinusoidal signals with frequencies from the set $\omega_{mn} = \pm m\omega_0 \pm n\omega_1$, where $m, n = 0, 1, 2, \cdots$:

$$\sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{mW_{mn}}{\omega_{mn}} = 0 \tag{1a}$$

$$\sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \frac{nW_{mn}}{\omega_{mn}} = 0. \tag{1b}$$

Here, W_{mn} is the average power entering the nonlinear reactance at ω_{mn}, and $W_{mn} = W_{-m,-n}$. The only restrictions on (1) are that ω_1/ω_0 must be irrational and that the nonlinear reactive element must have a single valued characteristic. For application to the problem of interest here, (1) reduces to

$$\frac{W_0 + W_{20}}{\omega_0} + \frac{W_2}{\omega_2} + \frac{W_3}{\omega_3} = 0 \tag{2a}$$

$$\frac{W_1}{\omega_1} + \frac{W_2}{\omega_2} - \frac{W_3}{\omega_3} = 0 \tag{2b}$$

where the double subscript notation has been dropped, except in the case of average power at the first pump harmonic, W_{20}.

It is important to note that the artifice of restricting the Manley-Rowe relations to a particular set of desired frequencies does not necessarily mean that a predicted result will be generally realizable without the presence of certain of the eliminated frequencies. For example, if the external circuitry reactively terminates a particular frequency, it will be eliminated from (1). However, through internal conversion this frequency may be a hidden mechanism behind effects that seem, from (1), to be independent of such frequency. This situation arises in this problem in connection with the first pump harmonic. Eq. (1a) suggests nothing about the relative importance of the pump and its harmonic, and, in fact, falsely suggests that they are interchangeable. Since it is the reactive power of the pump harmonic that is important in this case, W_{20} will be ignored for the present.

Further study of (2) will be divided into two parts. Consider first the application of signals at ω_1 and ω_0, with ω_2 and ω_3 seeing passive loads only (*i.e.*, loads whose resistive components are positive). Thus, W_2 and W_3 are nonpositive and (2b) yields the following expressions for up-conversion power gains from ω_1 to ω_2 or ω_3:

$$G_{p_{12}} = -\frac{W_2}{W_1} = \frac{\omega_2}{\omega_1} \frac{1}{1 - X_{32}} \tag{3a}$$

$$G_{p_{13}} = -\frac{W_3}{W_1} = \frac{\omega_3}{\omega_1} \frac{X_{32}}{1 - X_{32}} \tag{3b}$$

$$\frac{W_0}{W_1} = \frac{\omega_0}{\omega_1} \frac{1 + X_{32}}{1 - X_{32}} \tag{3c}$$

where $X_{32} = W_3\omega_2/W_2\omega_3$. To illustrate the remarks to follow, sketches of (3) are shown in Fig. 2. In the limit where ω_3 is short circuited ($W_3 = 0$), $G_{p_{12}}$ has the well-known value ω_2/ω_1. However, $G_{p_{12}}$ increases as X_{32} is increased, and arbitrarily-large positive values of $G_{p_{12}}$ can be achieved in the vicinity of $X_{32} \lesssim 1$. This suggests that large conversion gain is possible without reflecting negative input conductance at ω_1. For $X_{32} > 1$, $G_{p_{12}}$ and

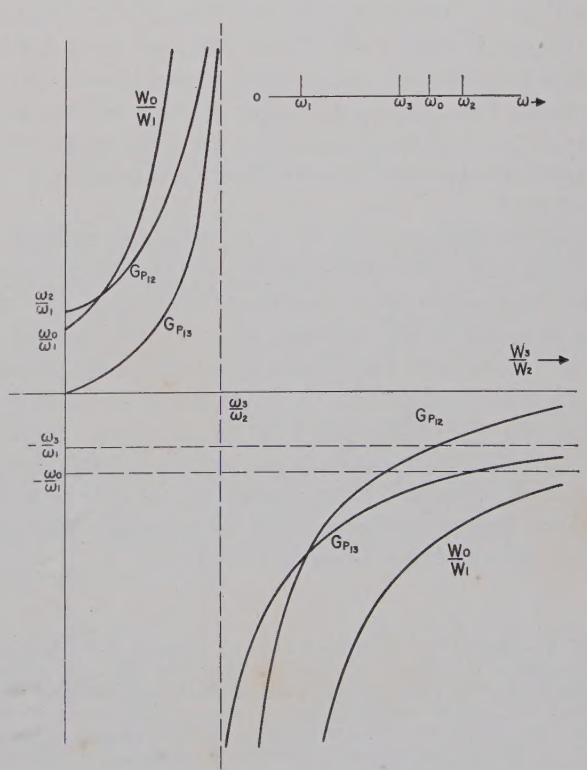


Fig. 2—Sketches of (3) showing the regions of positive and negative conversion gains from ω_1 to ω_2 and ω_3 vs the ratio of output powers at ω_3 and ω_2. Pump power is applied at ω_0, and W_0/W_1 is the ratio of pump to signal power.

$G_{p_{13}}$ are negative. Hence, this condition defines a general region of potential instability in which parametric amplification is possible at ω_1. It is important to note that negative conversion gains do not necessarily imply the reflection of negative conductance at the input frequency terminals, although this has been the case in previous nonlinear reactance circuits. Eq. (1) would also predict negative conversion gains if the input signal were completely absorbed at one terminal pair, amplified, and then expelled at the same frequency (along with other conversion frequencies) at another terminal pair. The realization of this effect (without using circulators) is one goal of this paper.

The second case of interest is with signals applied at ω_0 and ω_2, and with only passive loads for ω_1 and ω_3. In this case, (2) yields, with $X_{31} = W_3\omega_1/W_1\omega_3$,

$$G_{p_{21}} = -\frac{W_1}{W_2} = \frac{\omega_1}{\omega_2}\,\frac{1}{1 - X_{31}} \tag{4a}$$

$$G_{p_{23}} = -\frac{W_3}{W_2} = \frac{\omega_3}{\omega_2}\,\frac{X_{31}}{1 - X_{31}} \tag{4b}$$

$$\frac{W_0}{W_2} = \frac{\omega_0}{\omega_2}\,\frac{2X_{31} - 1}{1 - X_{31}} \tag{4c}$$

which are sketched in Fig. 3. For $X_{31} < 1/2$, G_{21} and G_{23} are positive, but W_0/W_2 is negative. Therefore, the pump circuit is unstable in this region and the signal at ω_2 will act as the power source. For $(\omega_0/\omega_2) < X_{31} < 1$, $G_{p_{21}}$ is greater than unity. Therefore, this is a region of arbitrary down-conversion gain with positive input conductance. For $X_{31} > 1$, $G_{p_{21}}$ and W_0/W_2 are negative. Thus, a region of potential instability exists at ω_2, which suggests that parametric amplification is possible at this frequency.[2]

Again it should be noted that (1) in no way guarantees that just any nonlinear reactance will yield the results above. For a particular nonlinear reactance, (1) may have only the trivial solution, $W_{mn} = 0$. For example, it can be noted from Fig. 3 that parametric amplification at ω_2 is predicted even in the limit where ω_1 is short circuited ($W_1 = 0$).[3] However, with the nonlinear reactance model employed in small-signal analysis (a model which satisfies the Manley-Rowe equations), it can be shown that the pump harmonic ($2\omega_0$) is necessary for parametric amplification at ω_2, if $W_1 = 0$.

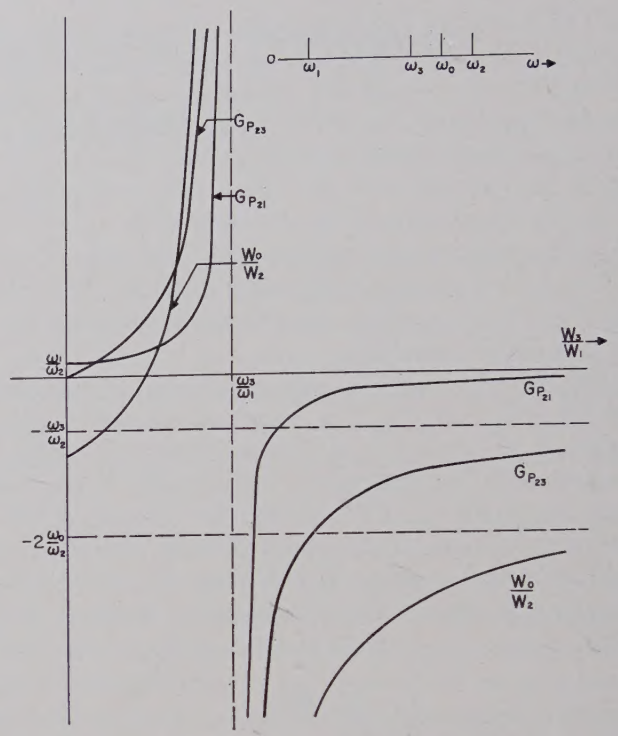


Fig. 3—Sketches of (4) showing the regions of positive and negative conversion gains fron ω_2 to ω_1 and ω_3 vs the ratio of output powers at ω_3 and ω_2. Pump power is applied at ω_0, and W_0/W_1 is the ratio of pump to signal power.

Small-Signal Theory

Although the Manley-Rowe equations provide a basis for predicting general properties of nonlinear reactance circuits, they yield no specific information on impedance levels, bandwidth, or the importance of extraneous frequencies. Such information requires further knowledge of the element in question plus mathematical techniques for handling the associated nonlinear circuits. Fortunately, several approximate characterizations of nonlinear elements lend themselves to standard analytical methods. In general, these involve some form of small-signal analysis, the most basic of which neglects responses of the nonlinear element to all signals except the power source and its harmonics. Small signals at all other frequencies then see essentially a periodically-varying circuit element at fundamental frequency ω_0.[4,5]

In this section, a nonlinear capacitor will be assumed to be driven by a power source at ω_0. To employ the small-signal model outlined above, the power source and nonlinear capacitor will immediately be replaced by a periodically-varying capacitor for which the following Fourier representation applies:

$$C(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_0 t} \tag{5}$$

[2] C. L. Hogan, R. L. Jepsen, and P. H. Vartanian, "New type of ferromagnetic amplifier," *J. Appl. Phys.*, vol. 29, pp. 422–423; March, 1958. This paper contains the original proposal of the possibility of parametric amplification ω_2, using ω_1 and ω_3 as so-called "idler" frequencies. Since the work seems to be based on purely qualitative arguments, parametric amplification at ω_2 will be treated in greater detail in the present paper.

[3] For previous treatment of this special case through small-signal analysis, see S. Bloom and K. K. N. Chang, "Parametric amplification using low-frequency pumping," Proc. IRE, vol. 46, pp. 1383–1387; July, 1958.

[4] H. E. Rowe, "Some general properties of nonlinear elements —Part II. Small signal theory," Proc. IRE, vol. 46, pp. 850–860; May, 1958.

[5] D. Lennov, "Gain and noise figure of a variable capacitance upconverter," *Bell Sys. Tech. J.*, vol. 37, pp. 989–1008; July, 1958.

where $C_n = C_{-n}^*$.[6] Now, assuming the external circuitry essentially short-circuits all small-signal frequencies except ω_1, ω_2, and ω_3, the small-signal voltage and current across the time-varying capacitor will be given by

$$V(t) = \sum_{n=1}^{3} (V_n e^{j\omega_n t} + V_n^* e^{-j\omega_n t}) \tag{6a}$$

$$I(t) = \sum_{n=1}^{3} (I_n e^{j\omega_n t} + I_n^* e^{-j\omega_n t}) + (\text{terms representing short circuit currents at other frequencies}). \tag{6b}$$

In (6b), it is necessary to consider only terms in ω_1, ω_2, and ω_3, since these are the only current components accessible to external terminals. By combining (5) and (6), the following admittance relation is found between the currents and voltages across the time-varying capacitor at each of the small-signal frequencies:[4]

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3^* \end{bmatrix} = \begin{bmatrix} j\omega_1 C_0 & j\omega_1 C_1^* & j\omega_1 C_1 \\ j\omega_2 C_1 & j\omega_2 C_0 & j\omega_2 C_2 \\ -j\omega_3 C_1^* & -j\omega_3 C_2^* & -j\omega_3 C_0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3^* \end{bmatrix}. \tag{7}$$

It is important to note that, because of the terms in C_2 in (7), small-signal effects on ω_1, ω_2, and ω_3 will generally include effects of both the pump and its first harmonic. Thus, for a more general formulation of this problem, terms in C_2 will be included.

Eq. (7) suggests the linear circuit in Fig. 4 as a small-signal model for this problem. Because of the assumption of ideal filters in Fig. 4, the external admittances are easily included with (7) to yield the complete admittance relationship shown below.

$$\begin{bmatrix} I_{10} \\ I_{20} \\ I_{30}^* \end{bmatrix} = \begin{bmatrix} Y_1 + j\omega_1 C_0 & j\omega_1 C_1^* & j\omega_1 C_1 \\ j\omega_2 C_1 & Y_2 + j\omega_2 C_0 & j\omega_2 C_2 \\ -j\omega_3 C_1^* & -j\omega_3 C_2^* & Y_3^* - j\omega_3 C_0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3^* \end{bmatrix}. \tag{8}$$

In general, two basic phenomena will be investigated with the aid of (8): 1) power conversion from ω_m to ω_n; and 2) parametric amplification at ω_m; where $m, n = 1, 2$ or 2, 1, respectively. Referring to Fig. 5, if $G_{t_{mn}}$ is the ratio of the output power at ω_n to the available generator power from a source at ω_m, then

$$G_{t_{mn}} = \frac{4g_s(g_{m_{IN}} - g_m)}{|Y_{m_{IN}}|^2} \times G_{p_{mn}} \tag{9}$$

where $Y_{m_{IN}} = g_{m_{IN}} + jb_{m_{IN}}$ is the driving point admittance at terminals (m), g_s is the internal admittance of the input generator, and $G_{p_{mn}}$ is given by (3a) or (4a).

[6] Asterisk (*) indicates complex conjugate.

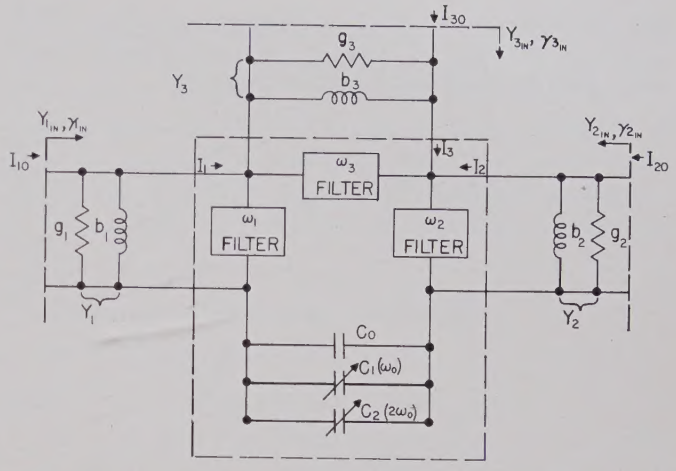


Fig. 4—A small-signal model, based on (7), for the nonlinear capacitance circuit shown in Fig. 1. Here, γ represents a normalized admittance, according to the definitions in (12).

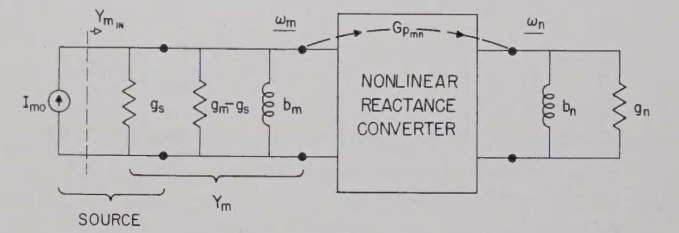


Fig. 5—The modified external admittance notation that is assumed whenever external sources are present, as in the derivation of (9).

Since $G_{p_{mn}}$ can be alternately expressed by

$$G_{p_{mn}} = \left|\frac{V_n}{V_m}\right|^2 \frac{g_n}{(g_{m_{IN}} - g_m)}, \tag{10}$$

(9) becomes

$$G_{t_{mn}} = \frac{4g_s g_n}{|Y_{m_{IN}}|^2} \times \left|\frac{V_n}{V_m}\right|^2. \tag{11}$$

Ultimately, conditions will be sought under which $G_{t_{mn}}$ is arbitrarily large. As will be shown shortly, this condition usually arises only when $|Y_{m_{IN}}|$ approaches zero. Since the vanishing of $|Y_{m_{IN}}|$ is also a sufficient condition for parametric amplification at ω_m, first attention will be directed toward this quanitity.

Since $Y_{1_{IN}} = (I_{10}/V_1)$ with $I_{20} = I_{30} = 0$, etc., the evaluation of $Y_{k_{IN}}$ $(k = 1, 2, \text{or } 3)$ from (8) is straightforward. Once obtained, however, the complexity of each $Y_{k_{IN}}$ expression can be considerably reduced by the introduction of normalized admittances γ_k and $\gamma_{k_{IN}}$, which are defined as follows:

$$\gamma_{1_{IN}} = \frac{Y_{1_{IN}}}{\omega_1}\left|\frac{C_2}{C_1^2}\right|; \qquad \gamma_1 = \frac{Y_1 + j\omega_1 C_0}{\omega_1}\left|\frac{C_2}{C_1^2}\right| \tag{12a}$$

$$\gamma_{2_{IN}} = \frac{Y_{2_{IN}}}{\omega_2 |C_2|}; \qquad \gamma_2 = \frac{Y_2 + j\omega_2 C_0}{\omega_2 |C_2|} \tag{12b}$$

$$\gamma_{3_{IN}} = \frac{Y_{3_{IN}}^*}{\omega_3 |C_2|}; \qquad \gamma_3 = \frac{Y_3^* - j\omega_3 C_0}{\omega_3 |C_2|}. \tag{12c}$$

The notation $\gamma_k = \alpha_k + j\beta_k$ and $\gamma_{k_{IN}} = \alpha_{k_{IN}} + j\beta_{k_{IN}}$ ($k = 1$, 2, and 3) will also be adopted. The input admittances now take the following normalized forms:

$$\gamma_{1_{IN}} = \gamma_1 + \frac{\gamma_3 - \gamma_2 + 2j\phi}{\gamma_2\gamma_3 - 1} \tag{13a}$$

$$\gamma_{2_{IN}} = \gamma_2 + \frac{\gamma_3 - \gamma_1 + 2j\phi}{\gamma_1\gamma_3 - 1} \tag{13b}$$

$$\gamma_{3_{IN}} = \gamma_3 - \frac{\gamma_1 + \gamma_2 - 2j\phi}{\gamma_1\gamma_2 + 1}. \tag{13c}$$

Here,

$$\phi = \frac{\mathrm{Re}(C_1^2 C_2^*)}{|C_1^2 C_2|} \tag{14}$$

and hence is a function of only the relative phases of C_1 and C_2. Eqs. (13) are assumed to apply simultaneously and express both the normalized input admittance presented to an ideal generator at any terminal, and the normalized output admittance presented to an ideal voltmeter at any terminal. Therefore, whenever real sources of meters are applied, their internal admittances will be assumed to be lumped with each Y_k or γ_k ($k = 1$, 2, 3), as in Fig. 5.

The form of each of the equations of (13) readily identifies the physical significance of each term. The first term in each expression for $\gamma_{k_{IN}}$ is the normalized load admittance at terminals k ($k = 1$, 2, 3), while the second term is the corresponding reflected admittance at terminals k. Inspection of (13) reveals that many sets of passive loads exist for which one or more of the $\gamma_{k_{IN}}$ have negative real parts. This situation will be interpreted as complete instability, and hence (13) will be considered physically significant only when the real parts of $\gamma_{1_{IN}}$, $\gamma_{2_{IN}}$, and $\gamma_{3_{IN}}$ are simultaneously positive. In other words, a negative reflected conductance at any terminal pair will be considered useful only when the external conductance at that terminal has the greater magnitude of the two.

Since circuit adjustments are desired that yield large values of $G_{t_{mn}}$ in (11), the zeros of $|Y_{m_{IN}}|$ and the poles of $|V_n/V_m|$ will be of immediate interest. It follows, from the inversion of (8), that the poles of $|V_n/V_m|$ occur when $|\gamma_n\gamma_3 - 1| = 0$ ($n = 1$, 2). However, by (12) and (13), this condition usually yields a pole for $|Y_{m_{IN}}|$ ($m = 1$, 2). When this is the case then, the zeros of $|Y_{m_{IN}}|$ (and hence $|\gamma_{m_{IN}}|$) are the only infinite gain points. The only exception to this observation occurs when (13a) and (13b) are made independent of γ_3, treated below, under Case 2.

Case 1. $|C_2| = 0$

The first application of (13) will be to the true four-frequency problem in which all effects of the pump harmonics are assumed negligible. This corresponds to letting $|C_2|$ approach zero, whereupon (13) reduces to the following limiting forms:

$$\gamma_{1_{IN}} = \gamma_1 + \frac{\gamma_3 - \gamma_2}{\gamma_2\gamma_3} \tag{15a}$$

$$\gamma_{2_{IN}} = \gamma_2 + \frac{\gamma_3}{\gamma_1\gamma_3 - 1} \tag{15b}$$

$$\gamma_{3_{IN}} = \gamma_3 - \frac{\gamma_2}{\gamma_1\gamma_2 + 1}. \tag{15c}$$

Note that $|C_2|$ cancels identically from each of the above expressions. Therefore, it will be required that the equations of (15), and those to follow in this case, have solutions other than the trivial one, $|C_2| = 0$.

Since (15a) and (15b) cannot be made independent of γ_3 in this case, only arbitrarily-small values of $|\gamma_{m_{IN}}|$ ($m = 1$ or 2) will allow unlimited conversion gains from ω_1 to ω_2, or vice versa. Therefore, when $|C_2| = 0$, the possible advantages of four-frequency circuits over conventional three-frequency circuits are: 1) the reflection of useful negative conductance at ω_1 in the presence of ω_2, and 2) the reflection of useful negative conductance at ω_2. Since the latter result is the more novel of the two, in that it provides parametric amplification at a frequency higher than the true pump frequency (*i.e.*, ω_0, since $|C_2| = 0$), primary consideration will be given to this result. This study will be done under two constraints: 1) $\alpha_{1_{IN}}$, $\alpha_{2_{IN}}$, and $\alpha_{3_{IN}}$ must be positive for stability, and 2) $\beta_{1_{IN}}$, $\beta_{2_{IN}}$, and $\beta_{3_{IN}}$ must be small (preferably zero) at the center operating frequency. The latter condition makes all three terminals nearly resonant, which is desirable as a practical means of eliminating the ideal filters that are tacit to this analysis. Only approximate resonance is required at the external terminals in this case, because it will be shown that some detuning is necessary for parametric amplification at ω_2 to occur.

For purposes of introduction, a special case of four-frequency operation will be described first. One technique for resonating the external terminals is to make $\beta_1 = \beta_2 = \beta_3 = 0$, which means physically that external reactances are chosen to resonate C_0 independently at each terminal. This choice of external reactance, which has been generally satisfactory in conventional three-frequency circuits, is attractive from the practical point of view because it allows circuit adjustments in the "cold" condition (*i.e.*, with the pump turned off). However, this tuning procedure will be shown in be relatively unproductive in this case.

Case 1(a). $|C_2|=0;\ \beta_1=\beta_2=\beta_3=0$

Under these conditions, (15) becomes

$$\alpha_{1_{IN}} = \alpha_1 + \frac{\alpha_3 - \alpha_2}{\alpha_3\alpha_2} \tag{16a}$$

$$\alpha_{2_{IN}} = \alpha_2 + \frac{\alpha_3}{\alpha_3\alpha_1 - 1} \tag{16b}$$

$$\alpha_{3_{IN}} = \alpha_3 - \frac{\alpha_2}{\alpha_1\alpha_2 + 1}. \tag{16c}$$

Inspection of (16) shows the condition for stability to be $\alpha_1\alpha_3>1$. Therefore, no useful region of negative reflected conductance exists at ω_2 and the minimum value of $\alpha_{1_{IN}}$ is limited by circuit stability to $1/\alpha_2$. Thus, parametric amplification at ω_2 is not possible in this case, and the presence of ω_2 limits the gain by parametric amplification at ω_1. However, some improvement in conversion gain from ω_1 to ω_2 is possible over that obtainable in the absence of ω_3 (*i.e.*, with $\alpha_3=\infty$). In this case, (11) yields $G_{t_{12}}=4\alpha_1\alpha_2\omega_2/\omega_1$ as the maximum up-conversion gain from ω_1 to ω_2. In the absence of ω_3, maximum $G_{t_{12}}$ occurs for $\alpha_1\alpha_2=1$. Therefore, allowing α_3 to take on finite values can increase the gain of a conventional up-converter by 6 db in this case, which is consistent with results reported by Jones and Honda.[7]

It will now be shown, through a more general investigation of (15), that the desired negative reflected conductances at ω_1 and ω_2 can be obtained with more arbitrary tuning techniques. These techniques are based on the following general theorems, which pertain to the four-frequency time-varying reactance circuit under consideration here.

Theorem 1: If $|C_2|=0$ and $\alpha_{1_{IN}},\ \alpha_{2_{IN}},\ \alpha_{3_{IN}}>0$, then either $\beta_{2_{IN}}\neq 0$, $\beta_{1_{IN}}\neq 0$, or $\beta_{3_{IN}}\neq 0$, is a necessary condition for parametric amplification at ω_2.

Proof: See Appendix.

Theorem 2: If $|C_2|=0$, the necessary and sufficient conditions for $\alpha_{1_{IN}},\ \alpha_{2_{IN}},\ \alpha_{3_{IN}}>0$ and $\beta_{1_{IN}}=\beta_{3_{IN}}=0$ (*i.e.*, a stable circuit with resonant terminals at ω_1 and ω_3) are:

$$\beta_3 = 0 \tag{17a}$$

$$\beta_2 = \beta_1|\gamma_2|^2 \tag{17b}$$

$$\alpha_1 - \frac{1}{\alpha_3} > \frac{-\alpha_2}{|\gamma_2|^2}\cdot\left[\text{lesser of}\left(1, \frac{\beta_2^2}{\alpha_2^2}\right)\right]. \tag{17c}$$

Proof: See Appendix.

It should be noted that Theorem 2 can be applied to conventional parametric amplifiers in the case where the upper sideband (ω_2) is not completely shorted, but where the pump harmonic is suppressed. In this case, (17b) indicates that "cold" tuning is not sufficient and some retuning must accompany changes in pump level.

In view of Theorem 1, parametric amplification at ω_2 will be approached by requiring only the terminals at ω_2 to be resonant ($\beta_{2_{IN}}=0$), for then $|Y_{2_{IN}}|$ approaches zero as $\alpha_{2_{IN}}$ approaches zero. The additional assumption that $\beta_3=0$ will be shown below to lead to useful solutions.

Case 1(b). $|C_2|=0;\ \beta_3=0;\ \beta_1\neq 0,\ \beta_2\neq 0$

Eq. (15) now yields

$$\frac{-\alpha_2}{|\gamma_2|^2} < \alpha_1 - \frac{1}{\alpha_3} < 0 \tag{18}$$

as the condition for: 1) stability at ω_1 and ω_3,[8] and 2) the reflection of negative conductance at ω_2. If (18) is satisfied, the condition for resonance at ω_2 then becomes

$$\left(\alpha_1 - \frac{1}{\alpha_3}\right)^2 + \left(\beta_1 - \frac{1}{2\beta_2}\right)^2 = \left(\frac{1}{2\beta_2}\right)^2. \tag{19}$$

If (19) is satisfied, the condition for stability at ω_2 is

$$\beta_2 < \beta_1|\gamma_2|^2. \tag{20}$$

These conditions are made clearer in Fig. 6, which is a plot of (19) for $\alpha_1\alpha_3<1$. Here, β_1 is considered to be the independent variable determining the magnitude of negative reflected conductance at ω_2. Point A is the threshold of parametric amplification at ω_2, point B marks the threshold of circuit instability at ω_1 [in (18)], while infinite gain at ω_2 [in (20)] corresponds either to point C, if $\beta_2^2<\alpha_2^2$, or to point B, if $\beta_2^2>\alpha_2^2$. Therefore, useful parametric amplification at ω_2 may be limited to the finite gains corresponding to the range between points A and B on the tuning curve in Fig. 6. This

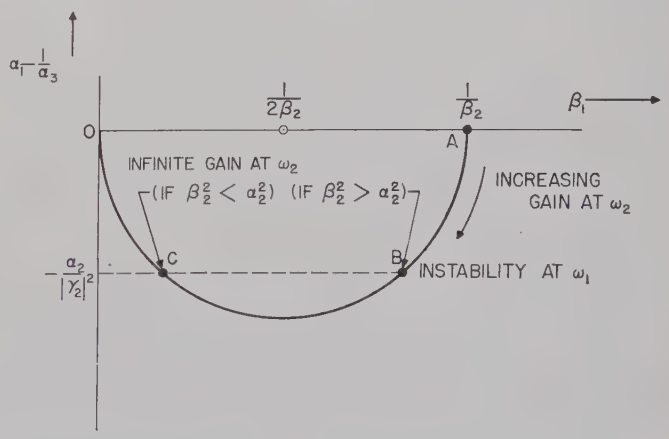


Fig. 6—The locus of external circuit adjustments for tuning and parametric amplification at ω_2, as described in Case 1(b).

[7] E. M. T. Jones and J. S. Honda, "A low-noise up-converter parametric amplifier," 1959 IRE WESCON Convention Record, pt. 1, pp. 99–103.

[8] If $\beta_3=0$, it can be shown, as α_3 decreases, that ω_3 never becomes unstable before ω_1 becomes unstable. Thus (18) only explicity defines the region of stability at ω_1.

limitation on gain can be overcome by choosing $\alpha_2^2 \leq \beta_2^2$. In this case, the negative conductances reflected at both ω_1 and ω_2 allow arbitrary gain at either frequency. It can be shown that a similar result is possible if the terminals at ω_1 are constrained to be resonant while those at ω_2 are detuned.

Thus, in the notation of Fig. 4 (but letting $b_k' = b_k + \omega_k C_0$, $k=1, 2, 3$), sufficient conditions for arbitrary gain at ω_2 (and at ω_1), with $|C_2| = 0$, are:

$$b_3' = 0 \tag{21a}$$

$$g_2^2 = K^2 b_2^2; \qquad K^2 \leq 1 \tag{21b}$$

$$\frac{1}{1+K^2} < \frac{b_1' b_2'}{\omega_1 \omega_2 |C_1|^2} < 1 \tag{21c}$$

$$\left(\frac{g_1}{\omega_1 |C_1|^2} - \frac{\omega_3}{g_3}\right)^2 + \left(\frac{b_1'}{\omega_1 |C_1|^2} - \frac{\omega_2}{2b_2'}\right)^2 = \left(\frac{\omega_2}{2b_2'}\right)^2. \tag{21d}$$

It is important to note that only conditions (20) or (21c) satisfy Theorem 2 at the point of infinite gain $[\beta_1\beta_2(1+K^2)=1]$. Thus, as previously mentioned, the terminals at ω_1 and ω_3 are not resonant in this case and suffer detuning at center frequency. For the special case of $K=1$, the detuning at ω_1 and ω_3 is given by

$$b_{1_{IN}} = b_1' - \frac{\omega_1\omega_2 |C_1|^2}{2b_2'} \tag{22a}$$

$$b_{3_{IN}} = \frac{2b_2'\omega_1\omega_3 |C_1|^2 (2b_1'b_2' - \omega_1\omega_2 |C_1|^2)}{(2g_1b_2' + \omega_1\omega_2 |C_1|^2)^2 + (2b_1'b_2' - \omega_1\omega_2 |C_1|^2)^2}. \tag{22b}$$

Eq. (22) indicates that the amount of detuning at ω_1 and ω_3 is very small when the gains at ω_2 (or at ω_1) are large.

Therefore, if $|C_2| = 0$, the operation described in Fig. 6 allows: 1) arbitrary gain by parametric amplification at either ω_1 or ω_2, and 2) arbitrary up- or down-conversion gain between ω_1 and ω_2. Both these effects arise through the reflection of negative conductance at the input terminals, and are accompanied by some detuning of the external terminals.

A physical interpretation of these results is suggested by (8). If $C_2=0$, it can be seen that ω_2 and ω_3 are not directly coupled, but interact only through the intermediary frequency ω_1. When considered separately, the feedback between ω_1 and ω_2 is degenerative, while that between ω_1 and ω_3 is regenerative. The combination of these two feedback mechanisms can make the over-all feedback at ω_2 regenerative. However, it has been found that additional phase shift must be inserted into these feedback paths (*e.g.*, by detuning the terminals at ω_1 and ω_3) to stabilize this regenerative effect.

Case 2: $|C_2| \neq 0$

It has been shown previously that the design of idealized, three- or four-frequency, nonlinear reactance converters for large, small-signal, conversion gains (*i.e.*, small-signal gains considerably greater than the ratio of output frequency to input frequency) must invariably rely on the reflection of negative resistance at the input terminals. Consequently, such devices frequently employ circulators as auxiliary equipment, in an effort to reduce noise figures and to increase gain-bandwidth products. It will now be shown, when $|C_2| \neq 0$, that arbitrarily-large conversion gains are possible between ω_1 and ω_2, without reflecting negative resistance at either frequency. Thus the importance of circulators in high-gain nonlinear reactance converters can be eliminated.

The technique for achieving the result cited above is to choose external circuit components that allow $|V_n/V_m|$ to be arbitrarily large [see (11)] without producing a similar effect on $Y_{m_{IN}}$ ($m, n = 1, 2$ or $2, 1$, respectively). It has been remarked previously that this condition can be achieved by making (13a) and (13b) independent of γ_3, which, in turn, can be accomplished only if

$$\gamma_1 = \gamma_2 = \sqrt{1-\phi^2} + j\phi. \tag{23}$$

Obviously, only the positive square root is permitted here, as will be the case henceforth, unless $(\pm)$ signs are specifically indicated. Substitution of (23) into (13a) and (13b) yields

$$\gamma_{1_{IN}} = \gamma_{2_{IN}} = 2\sqrt{1-\phi^2}. \tag{24a}$$

Thus, when (23) is satisfied, the terminals at ω_1 and ω_2 are conjugately matched for all γ_3. Similarly, (13c) reduces to

$$\gamma_{3_{IN}} = \gamma_3 - \sqrt{1-\phi^2} + j\phi \tag{24b}$$

which indicates that a constant negative conductance is reflected at ω_3. Consequently, $\alpha_3 > \sqrt{1-\phi^2}$ must be required for stability.

Although $\gamma_{m_{IN}}$ ($m=1, 2$) has been made independent of γ_3, it can be shown that $G_{t_{mn}}$ still has a useful pole as a function of γ_3. By solving (8) for $|V_n/V_m|$, and using (12), (11) yields the following expressions ($g_m = g_s$):

$$G_{t_{12}} = \frac{\omega_2}{\omega_1}\left[\frac{(\alpha_3 \pm \sqrt{1-\phi^2})^2 + (\beta_3+\phi)^2}{(\alpha_3 - \sqrt{1-\phi^2})^2 + (\beta_3+\phi)^2}\right] \tag{25a}$$

$$G_{t_{21}} = \frac{\omega_1}{\omega_2}\left[\frac{(\alpha_3 \mp \sqrt{1-\phi^2})^2 + (\beta_3+\phi)^2}{(\alpha_3 - \sqrt{1-\phi^2})^2 + (\beta_3+\phi)^2}\right]. \tag{25b}$$

Here, the $(\pm)$ signs arise because ϕ specifies only the magnitude of the relative phase angle between C_1 and C_2, and not its sign. In each equation of (25), the upper sign corresponds to Im $(C_1^2C_2^*)>0$ while the lower sign corresponds to Im $(C_1^2C_2^*)<0$. Thus, for a given value of ϕ, there are two possible up-conversion gains and two possible down-conversion gains. In the case of negative square roots, both $G_{t_{12}}$ and $G_{t_{21}}$ reduce to the ratio of output to input frequency, independently of γ_3; while with the positive square roots both $G_{t_{12}}$ and $G_{t_{21}}$ become arbitrarily large as γ_3 approaches the value $\sqrt{1-\phi^2}-j\phi$ (assuming, for stability, that it does so with $\alpha_3>\sqrt{1-\phi^2}$). Within a single converter stage, however, (25) cannot be satisfied simultaneously with positive square roots, which yields the useful result that each converter stage is semidirectional. The directivity is dependent on both the forward gain and the relative magnitudes of the frequencies employed.

Perhaps the most practical operating condition is with the terminals at γ_3 resonant (*i.e.*, $\beta_3=-\phi$), whereupon (25) reduces to

$$G_{t_{12}}=\frac{\omega_2}{\omega_1}\left[\frac{\alpha_3\pm\sqrt{1-\phi^2}}{\alpha_3-\sqrt{1-\phi^2}}\right]^2 \tag{26a}$$

$$G_{t_{21}}=\frac{\omega_1}{\omega_2}\left[\frac{\alpha_3\mp\sqrt{1-\phi^2}}{\alpha_3-\sqrt{1-\phi^2}}\right]^2. \tag{26b}$$

In the positive square root cases, the conversion gains now depend entirely on α_3, which means physically that variation of the external circuit Q at ω_3 controls conversion gain without changing the input or output admittance at any terminal pair.

The fact that two completely different conversion characteristics arise according to the sign of Im $(C_1^2C_2^*)$ is somewhat surprising, since the input and output admittances at all three terminals are independent of this sign. From Figs. 2 and 3 it is evident that the fixed gain cases of (25) and (26) must correspond to $W_3=0$, which can be verified by solving (8) for $|V_3|$. With sources first at ω_1 and then at ω_2, $|V_3|$ is found to vanish independently of γ_3 whenever the negative square roots occur in $G_{t_{12}}$ or $G_{t_{21}}$, respectively. However, in the positive square root cases, $|V_3|$ increases with $G_{t_{12}}$ and $G_{t_{21}}$, for sources at ω_1 and ω_2, respectively, which is again in agreement with Figs. 2 and 3.

It is interesting to assume also a source at ω_3 and to evaluate the quantities $|V_1/V_3|$ and $|V_2/V_3|$. It can be shown that ω_3 couples only to the ω_2 terminals when $G_{t_{12}}$ employs the positive square root, and that ω_3 couples only to the ω_2 terminals with the positive square root in $G_{t_{21}}$. Thus, the coupling between ω_3 and ω_1 or ω_2 is nonreciprocal in this case, and its direction depends on the sign of Im$(C_1^2C_2^*)$.

In summary, for a given magnitude of phase angle between the pump and its harmonics, the above results require the following external circuit adjustments:

At ω_1:

$$b_1+\omega_1C_0=\phi;\qquad Q_1=\frac{C_0}{\sqrt{1-\phi^2}}\left|\frac{C_2}{C_1^2}\right| \tag{27a}$$

At ω_2:

$$b_2+\omega_2C_0=\phi;\qquad Q_2=\frac{C_0}{|C_2|\sqrt{1-\phi^2}} \tag{27b}$$

At ω_3:

$$b_3+\omega_3C_0=\phi;\qquad Q_3<Q_2. \tag{27c}$$

It can be seen from (27) that each external terminal will be detuned in the absence of the pump, unless $\phi=0$. Also, the necessary external Q's increase as ϕ approaches unity. Therefore, the condition $\phi=0$ appears to be the preferable operating condition, since it calls for minimum Q circuits and permits circuit tuning prior to the application of the pump signals.

Since the results of this section have led to a small-signal converter with: 1) arbitrary power gain, and 2) conjugately matched input and output terminals, several applications immediately suggest themselves. The first application, which is shown in Fig. 7, is to cascaded up-converters. Here, the gain of each stage can be controlled individually at the respective lower sidebands, without requiring any adjustments of the interstage networks. Since each stage of this cascade is an up-converter, it is necessary that Im $(C_1^2C_1^*)>0$ for each stage, which makes the maximum gain in the reverse direction the ratio of input frequency to output frequency. A similar result follows if the sign of Im $(C_1^2C_2^*)$ is reversed in each stage, to make the cascade a high gain down-converter.

A modification of this application is shown in Fig. 8, where an up-converter and a down-converter are placed "back to back," such that the input and output frequencies are the same. In this configuration, the only basic difference between the two stages is in the sign of Im $(C_1^2C_2^*)$. Therefore, this circuit yields a nonlinear reactance amplifier with: 1) positive input and output conductance, 2) arbitrary forward gain, and 3) a maximum reverse gain of unity.

It is interesting to apply the Manley-Rowe equations to the circuit in Fig. 8. Since only ω_1 and ω_3 appear at the external terminals, this circuit is described by the same special case of (1) as is the conventional nonlinear reactance amplifier, where [by (3b), with $W_2=0$]

$$G_{p_{13}}=-\frac{W_3}{W_1}=-\frac{\omega_3}{\omega_1}. \tag{28}$$

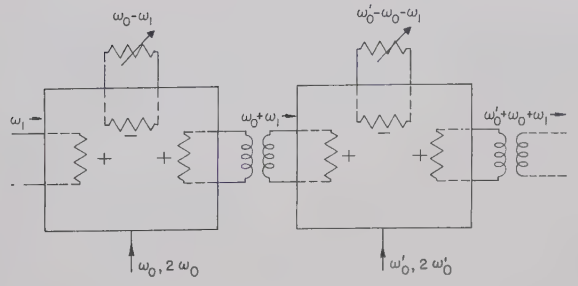


Fig. 7—A cascade of upper-sideband up-converters of the type described in Case 2. The reflected conductances, which are indicated in each box with appropriate signs, are gain-insensitive. Each stage is theoretically capable of unlimited gain.

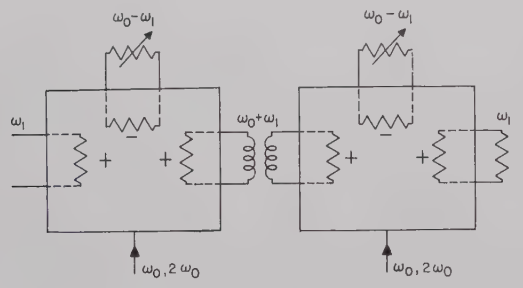


Fig. 8—A positive-input-conductance, nonlinear reactance amplifier formed by cascading an up-converter and a down-converter of the type described in Case 2. The only basic difference between these two stages is in the sign of the phase angle between ω_0 and $2\omega_0$.

In this case, however, W_1 has components at two separate terminals and hence can be negative even though the input is conjugately matched.

Although the circuits described above may have desirable gain and control aspects, their noise characteristics will serve as the ultimate test of their value. The primary sources of noise are associated with the external resistances and with internal resistance that is inherently present in any real nonlinear reactance element. In practice, the noise temperature of the output load will generally be fixed and may be large, which is one reason for the use of circulators in conventional negative-input-resistance amplifiers and converters. However, with the possibility for conjugate match offered in this case, output noise that is fed back to the input (because of the semidirectional nature of the circuits under consideration) can largely be absorbed. In addition, the "dummy load" nature of g_3 allows its temperature to be varied independently of the input and output circuitry. Thus, the so-called idler noise that is present in conventional nonlinear reactance circuits should be reduced considerably in this circuit. It can be proposed, therefore, even without the use of circulators, that the noise figure of this circuit should be improved over that of conventional negative-input-resistance amplifiers and converters, even when these latter circuits employ circulators.

Conclusion

A detailed analysis of (8), or equivalently, Fig. 4, has demonstrated several useful extensions of conventional nonlinear reactance circuits. It has been shown that parametric amplification at a frequency higher than the pump can, from the point of view of the nonlinear reactance element, be the same order of effect as conventional parametric amplification at a frequency lower than the pump. Also, a nonlinear-reactance circuit has been developed that is capable of unlimited up- or down-conversion gain, without reflecting negative resistance at either the input or the output terminal. The latter result suggests a means by which a nonlinear-reactance amplifier can be obtained which does not require a circulator for optimum performance. A more detailed summary of the results of this analysis is given in Table I. Within the limits of approximation provided by the circuit model in Fig. 4, Table I contains a compilation of the more important conclusions regarding three- and four-frequency, nonlinear reactance circuits.[9]

Appendix

Proofs of Theorems 1 and 2 can be obtained by the following considerations. These theorems deal with the real parts ($\alpha_{k_{IN}}$) and the imaginary parts ($\beta_{k_{IN}}$) of $\gamma_{k_{IN}}$ ($k=1$, 2, or 3), and explicit expressions for these quantities can be obtained from (15). The proof of Theorem 1 will be divided into the following steps.

Step A. There are two possible conditions under which $\beta_{1_{IN}}=\beta_{3_{IN}}=0$ can be satisfied simultaneously; namely,

$$\beta_3 = 0 \tag{29a}$$

or

$$\left|\frac{\gamma_3}{\gamma_2}\right|^2 = \frac{1}{|\gamma_1\gamma_2+1|^2}. \tag{29b}$$

However, for $\alpha_{1_{IN}}>0$ and $\alpha_{3_{IN}}>0$ to be satisfied simultaneously, it is necessary that

$$\frac{1}{|\gamma_1\gamma_2+1|^2} < \frac{\alpha_3}{\alpha_2+\alpha_1|\gamma_2|^2} < \left|\frac{\gamma_3}{\gamma_2}\right|^2. \tag{30}$$

Therefore, of the two possible conditions described, only (29a) will yield both stability and resonance at ω_1 and ω_3.

Step B. If $\beta_3=0$, the condition for resonance at ω_1 and ω_3 becomes $\beta_1|\gamma_2|^2=\beta_2$. Now, if $\beta_{2_{IN}}=0$ is also required, the following expression for $\alpha_{2_{IN}}$ can be obtained.

$$\alpha_{2_{IN}} = \alpha_2 \pm \alpha_2\left(1+\frac{\beta_2\beta_{2_{IN}}}{\alpha_2^2}\right)^{1/2}\left(1-\frac{\beta_{2_{IN}}}{\beta_2}\right)^{1/2} > 0. \tag{31}$$

[9] The term "four-frequency" has been employed throughout this paper even though a fifth frequency, in the form of the pump harmonic, has frequently been included. This choice of nomenclature has been made because only one of the two pump components need be applied externally, and hence only four-frequencies need be observed at external terminals.

TABLE I

A Summary of the Major Effects Obtainable from the Nonlinear Reactance Circuit Model of Fig. 4*

Frequency	Pump Components		
	(a) C_1 only	(b) C_2 only	(c) C_1 and C_2
(1) ω_1, ω_2	C.G.: Limited P.A.: None I.R.: Positive T: Gain independent	No effects	Same as (1a)
(2) ω_1, ω_3	C.G.: Unlimited P.A.: Unlimited I.R.: Negative T: Gain independent	No effects	Same as (2a)
(3) ω_2, ω_3	No effects	Same as (2a)	Same as (2a)
(4) $\omega_1, \omega_2, \omega_3$	C.G.: Unlimited P.A.: Unlimited I.R.: Negative T: Gain-dependent; Tuning impossible at all three frequencies	Same as (3b)	C.G.: Unlimited P.A.: Unlimited I.R.: Positive at ω_1 and ω_2 Negative at ω_3 T: Gain independent

* In entries where one or more frequencies or time-varying capacity components are omitted, they are assumed to be short-circuited. Key: C.G. = conversion gain; P.A. = gain by parametric amplification; I.R. = input resistance; and T = tuning conditions.

However, for parametric amplification at ω_2, (31) must be satisfied in such a way that the second term takes on the negative sign. Clearly, this is not possible if $\beta_{2_{IN}}=0$. Consequently, if $\beta_{1_{IN}}=\beta_{3_{IN}}=0$, then $\beta_{2_{IN}}\neq 0$ is necessary for parametric amplification at ω_2. $\beta_{2_{IN}}=0$ can occur only if $\alpha_{2_{IN}}$ equals $2\alpha_2$ or zero (*i.e.*, only under conditions of conjugate match or infinite gain at ω_2).

Step C. This step amounts simply to noting that, under Case 1(b) of the text, $\beta_{2_{IN}}=0$ and $\beta_3=0$ are shown to yield $\beta_{1_{IN}}\neq 0$ and $\beta_{3_{IN}}\neq 0$ in the region of parametric amplification at ω_2.

Therefore, under the conditions stated in Theorem 1, $\beta_3=0$ yields either $\beta_{1_{IN}}=\beta_{3_{IN}}=0$ and $\beta_{2_{IN}}\neq 0$ (Step B), or $\beta_{2_{IN}}=0$, $\beta_{1_{IN}}\neq 0$, and $\beta_{3_{IN}}\neq 0$ (Step C). Similarly, with $\beta_3\neq 0$, $\beta_{1_{IN}}$ and $\beta_{3_{IN}}$ cannot vanish simultaneously (Step A). Therefore, stable parametric amplification at ω_2, with $|C_2|=0$, requires at least one of the three quantities, $\beta_{1_{IN}}$, $\beta_{2_{IN}}$, or $\beta_{3_{IN}}$, not to vanish, which proves Theorem 1.

Now consider Theorem 2. By Steps A and B above, the necessary and sufficient conditions for ω_1 and ω_3 to be both resonant and stable[8] are:

$$\beta_3 = 0 \tag{32a}$$

$$\beta_2 = \beta_1 |\gamma_2|^2 \tag{32b}$$

$$\alpha_1 - \frac{1}{\alpha_3} > \frac{-\alpha_2}{|\gamma_2|^2}. \tag{32c}$$

If (32a) is satisfied,

$$\alpha_1 - \frac{1}{\alpha_3} > \frac{-1}{2\alpha_2}\left[1-(1-4\alpha_2^2\beta_1^2)^{1/2}\right] \tag{33}$$

is necessary and sufficient for stability at ω_2. However, by (32b), (33) can be written

$$\alpha_1 - \frac{1}{\alpha_3} > -\frac{\alpha_2}{|\gamma_2|^2}\left[\text{lesser of}\left(1, \frac{\beta_2^2}{\alpha_2^2}\right)\right] \tag{34}$$

which is the necessary and sufficient condition for stability at ω_2. Since (34) is also sufficient for stability at ω_1 and ω_3, the proof of Theorem 2 is completed.

Acknowledgment

The author wishes to express his appreciation to Professor C. B. Sharpe and to R. T. Denton for frequent helpful discussions during the course of this research.

Some Properties of Three Coupled Waves*

LASZLO SOLYMAR†

Summary—The paper deals with the problem of three waves, 1, 2, and 3, in which waves 2 and 3 are coupled to wave 1 but not to each other. The general solution for the amplitudes of the waves is given in closed form. It is shown that for certain values of the parameters growing waves can exist. Numerical solutions for the location of the boundaries of the growing wave regions are plotted. It is shown furthermore that under certain conditions the power can be completely transferred from wave 1 to waves 2 and 3.

Examples on traveling-wave tubes, waveguide couplers, and backward-wave oscillators illustrate the applicability of the theory.

I. Introduction

COUPLED wave theory[1–3] has in the recent past proved to be a powerful approach to the approximate solution of a wide variety of problems. Not only are the results often quantitatively of sufficient accuracy, but the physical picture which emerges is also of great value in understanding the essential nature of the particular problem.

This paper is an attempt to extend the quantitative treatment to three lossless coupled waves. In many cases it turns out that two of the three waves are uncoupled. Accordingly, this restriction has been imposed in the paper, with a considerable saving in complexity. This picture can be successfully applied to the description of waveguide couplers, traveling-wave tubes, and backward-wave oscillators, but naturally the conclusions are much more general and are valid for any coupled system.

In Section II the general solution of the coupled wave differential equation system is given in closed form. In Section III the condition for growing waves is found and the results are plotted in Figs. 1–8. In Section IV conditions of complete power transfer are investigated. In Section V four examples are given which demonstrate the applicability of the general formulas derived.

II. The Solution of the Coupled Wave Differential Equation System

The generality of the solution will be restricted in the following aspects:

1) Wave 2 and wave 3 are not coupled.
2) The couplings between waves 1 and 2, and waves 1 and 3 are assumed to be uniform, *i.e.*, they are independent of the space variable z.
3) The phase velocities of all three waves are in the positive direction of the z axis.
4) At the beginning of the coupled system all the power is in wave 1.

Subject to the above restrictions the coupled wave differential equation system can be written as follows:[4]

$$\begin{aligned} -\frac{dE_1}{dz} &= j\beta_1 E_1 + jd_{12}E_2 + jd_{13}E_3 \\ -\frac{dE_2}{dz} &= jf_{12}d_{12}E_1 + j\beta_2 E_2 \\ -\frac{dE_3}{dz} &= jf_{13}d_{13}E_1 + j\beta_3 E_3 \end{aligned} \tag{1}$$

where

E_1, E_2, E_3 = the amplitudes of waves 1, 2, and 3 respectively,
$\beta_1, \beta_2, \beta_3$ = the propagation coefficients of waves 1, 2, and 3 respectively,
d_{12}, d_{13} = coupling coefficients between waves 1 and 2, and 1 and 3 respectively, and
f_{12}, f_{13} = ± 1 if the energy velocity of wave 2, 3 is in the same/opposite direction as that of wave 1.

We now solve the differential equation system by assuming the following form for the amplitudes:

$$\begin{aligned} E_1 &= A_1 e^{jt_1 z} + A_2 e^{jt_2 z} + A_3 e^{jt_3 z} \\ E_2 &= B_1 e^{jt_1 z} + B_2 e^{jt_2 z} + B_3 e^{jt_3 z} \\ E_3 &= C_1 e^{jt_1 z} + C_2 e^{jt_2 z} + C_3 e^{jt_3 z}. \end{aligned} \tag{2}$$

Substituting (2) into (1) the unknown coefficients can be determined, while t_1, t_2, t_3 are the roots of the following third power equation:

$$t^3 + T_1 t^2 + T_2 t + T_3 = 0 \tag{3}$$

where

$$\begin{aligned} T_1 &= \beta_1 + \beta_2 + \beta_3 \\ T_2 &= \beta_1\beta_2 + \beta_1\beta_3 + \beta_2\beta_3 - f_{12}d_{12}^2 - f_{13}d_{13}^2 \\ T_3 &= \beta_1\beta_2\beta_3 - f_{12}d_{12}^2\beta_3 - f_{13}d_{13}^2\beta_2. \end{aligned} \tag{4}$$

Applying furthermore the boundary conditions, in accordance with restriction 4),

$$E_1(0) = 1, \qquad E_2(0) = 0, \qquad E_3(0) = 0. \tag{5}$$

* Manuscript received by the PGMTT, September 29, 1959.
† Standard Telecommunications Labs., Ltd., Harlow, Eng.
[1] J. R. Pierce, "Coupling of modes of propagation," *J. Appl. Phys.*, vol. 25, pp. 179–183; February, 1954.
[2] S. E. Miller, "Coupled wave theory and waveguide applications," *Bell. Sys. Tech., J.*, vol. 33, pp. 661–719; May, 1954.
[3] J. R. Pierce, "The wave picture of microwave tubes," *Bell Sys. Tech. J.*, vol. 33, pp. 1343–1372; November, 1954.
[4] The relations between the matrix elements are a direct consequence of the conservation of energy.

The general solution can be written in the following closed form:

$$E_1(z) = \sum_{i=1}^{3} \frac{(t_i+\beta_2)(t_i+\beta_3)}{(t_i-t_{i+1})(t_i-t_{i-1})} e^{jt_iz}$$

$$E_2(z) = -f_{12}d_{12} \sum_{i=1}^{3} \frac{t_i+\beta_3}{(t_i-t_{i+1})(t_i-t_{i-1})} e^{jt_iz}$$

$$E_3(z) = -f_{13}d_{13} \sum_{i=1}^{3} \frac{t_i+\beta_2}{(t_i-t_{i+1})(t_i-t_{i-1})} e^{jt_iz} \quad (6)$$

where

$$t_0 = t_3 \quad \text{and} \quad t_4 = t_1.$$

III. The Condition for the Existence of Complex Roots

The solution of (3) may result in three real roots or in one real and two complex roots. It may be seen from (6) that complex roots mean an attenuating and a growing wave. It should be appreciated, however, that the existence of a growing wave solution does not necessarily imply "amplification" in the usual sense. This will always depend on the boundary conditions imposed by the physics of the problem. In fact, amplification can take place even when all the roots are purely imaginary, two examples being the backward-wave amplifier and the crestatron. Nevertheless, in most physical problems, the demarcation between the regions of pure imaginary and complex roots is of fundamental significance.

By introducing the new variable

$$u = t + \tfrac{1}{3}T_1 \quad (7)$$

we bring (3) into the following more appropriate form:

$$u^3 + 3Hu + G = 0 \quad (8)$$

where

$$3H = T_2 - \tfrac{1}{3}T_1^2$$
$$= -\tfrac{1}{6}[r^2 + p^2 + (r-p)^2 + 6(f_{12}d_{12}^2 + f_{13}d_{13}^2)] \quad (9)$$

$$G = \frac{2}{27}T_1^3 - \tfrac{1}{3}T_1T_2 + T_3 = -\frac{1}{27}[(p+r)(2p-r)$$
$$\cdot(2r-p) + 9(2p-r)f_{12}d_{12}^2 + (2r-p)f_{13}d_{13}^2] \quad (10)$$

$$p = \beta_3 - \beta_1, \qquad r = \beta_2 - \beta_1. \quad (11)$$

It can be seen that both H and G depend only on the difference of the propagation coefficients. This is physically obvious, because it is always possible to regard one of the waves as stationary.

Now we can express in mathematical form the condition for complex roots. Eq. (3) has two complex roots, if

$$M = G^2 + 4H^3 > 0. \quad (12)$$

Substituting (9) and (10) into (12) and arranging by powers of p, we obtain:

$$M = -p^4(r^2 + 4f_{12}d_{12}^2) + 2p^3r(r^2 - f_{13}d_{13}^2 + 4f_{12}d_{12}^2)$$
$$- p^2[r^4 + 2r^2(f_{12}d_{12}^2 + f_{13}d_{13}^2) - 8d_{12}^4$$
$$+ 20f_{12}f_{13}d_{12}^2d_{13}^2 + d_{13}^4]$$
$$+ 2pr[r^2(4f_{13}d_{13}^2 - f_{12}d_{12}^2) - 4d_{12}^4$$
$$+ 19f_{12}f_{13}d_{12}^2d_{13}^2 - 4d_{13}^4]$$
$$- 4r^4f_{13}d_{13}^2 - r^2(d_{12}^4 + 20f_{12}f_{13}d_{12}^2d_{13}^2 - 8d_{13}^4)$$
$$- 4(f_{12}d_{12}^2 + f_{13}d_{13}^2)^3. \quad (13)$$

Since M depends on the direction of the energy velocities, v_{e1}, v_{e2}, v_{e3}, we have to investigate three cases. Denoting an energy velocity in the same direction as wave 1 by s, and in the opposite direction as wave 1 by o, we have the following three cases:

	(a)	(b)	(c)
v_{e2}	s	s	o
v_{e3}	s	o	o.

The fourth possibility (os) has been omitted as we are not distinguishing between waves 2 and 3. It can be shown from (13) that case (a) always leads to $M<0$ so that here no growing wave solution exists.

The study of case (b) reveals (Figs. 1–5) that for certain values of p, r, d_{12}, d_{13} growing wave solution exists. The figures show the $M=0$ lines on the p/d_{12}, r/d_{12} plane for different values of d_{13}^2/d_{12}^2. The curves are plotted only for positive values of p/d_{12}, because of the relation $M(p/d_{12}, r/d_{12}) = M(-p/d_{12}, -r/d_{12})$. Each of these Figures can be roughly divided into three parts:

1) The neighborhood of the origin,
2) the neighborhood of the r/d_{12} axis (except near the origin), and
3) the neighborhood of the $p=r$ line (except near the origin).

The following conclusions can be drawn for each, respectively:

1) The greater the coupling to wave 3 compared with that to wave 2, the greater is the extent of the growing wave region near the origin.
2) If $p=\beta_3-\beta_1\approx 0$, *i.e.*, the velocity of wave 3 is near to that of wave 1, growing wave solution always exists irrespective of the value of r/d_{12}. As r/d_{12} goes to $(\pm\infty)$ the $M=0$ line approaches the $p/d_{12}=2d_{13}/d_{12}$ asymptote (broken lines).
3) If $p>r>0$, but p and r are nearly equal, growing wave solution exists. This means physically, that if wave 3 is slower than wave 2 (both being slower than wave 1), but the velocity difference between waves 2 and 3 is sufficiently small, growing wave solution exists irrespective of the velocity of wave 1.

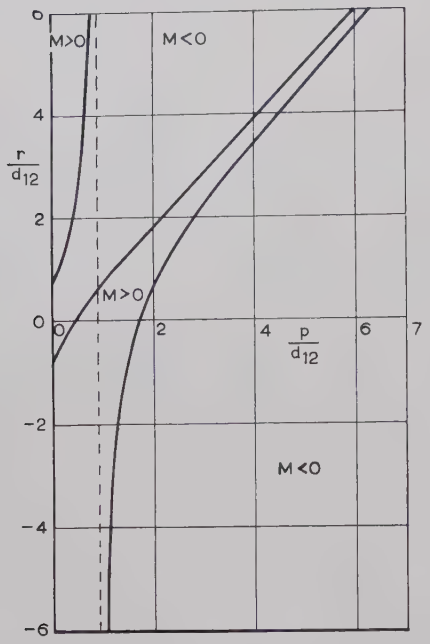


Fig. 1—The domain of complex roots for $f_{12}=1, f_{13}=-1, d_{13}^2/d_{12}^2=1/5$.

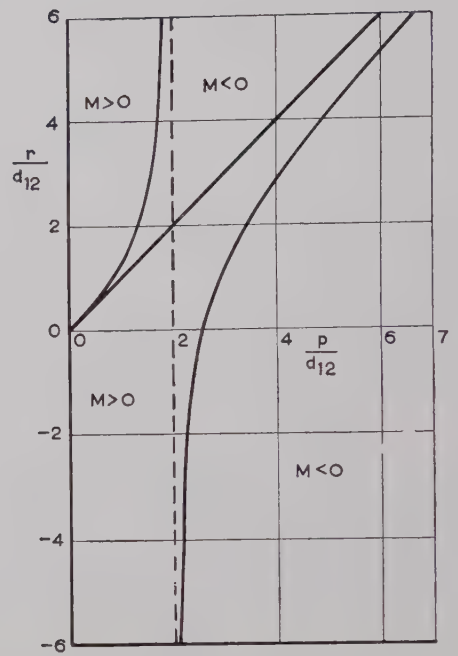


Fig. 3—The domain of complex roots for $f_{12}=1, f_{13}=-1, d_{13}^2/d_{12}^2=1$.

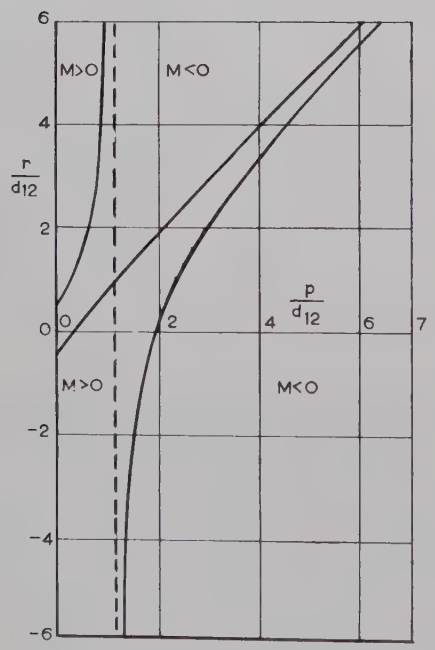


Fig. 2—The domain of complex roots for $f_{12}=1, f_{13}=-1, d_{13}^2/d_{12}^2=1/3$.

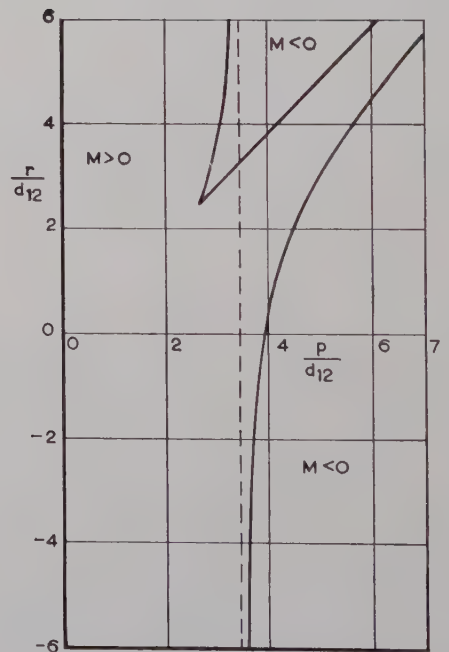


Fig. 4—The domain of complex roots for $f_{12}=1, f_{13}=1, d_{13}^2/d_{12}^2=3$.

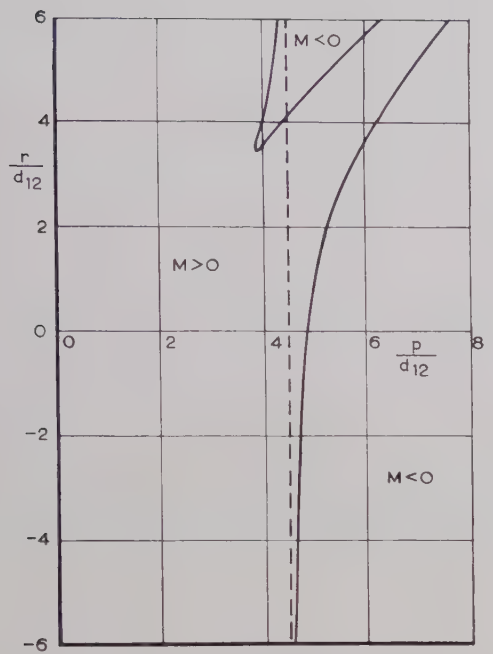


Fig. 5—The domain of complex roots for $f_{12}=1$, $f_{13}=-1$, $d_{13}^2/d_{12}^2=5$.

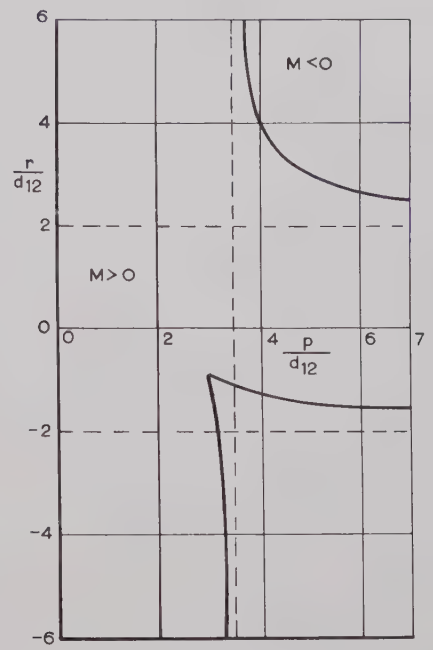


Fig. 7—The domain of complex roots for $f_{12}=-1$, $f_{13}=-1$, $d_{13}^2/d_{12}^2=3$.

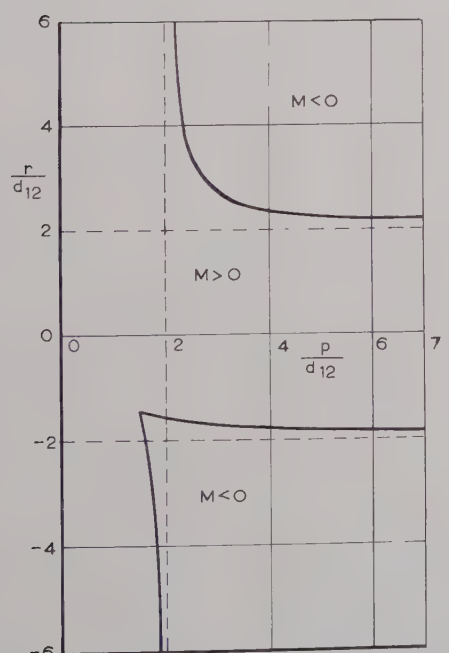


Fig. 6—The domain of complex roots for $f_{12}=-1$, $f_{13}=-1$, $d_{13}^2/d_{12}^2=1$.

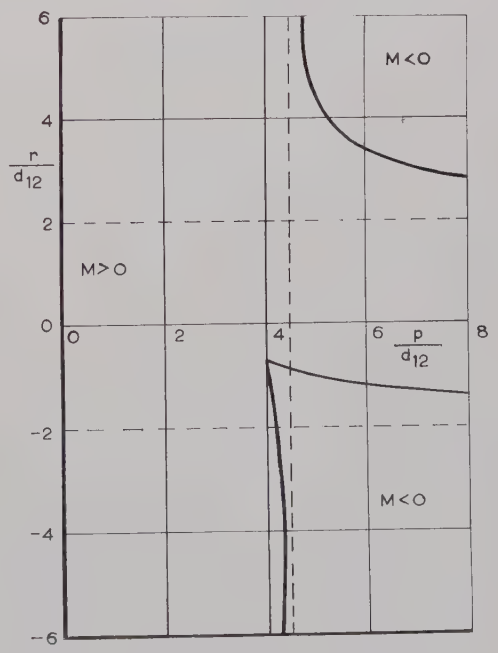


Fig. 8—The domain of complex roots for $f_{12}=-1$, $f_{13}=-1$, $d_{13}^2/d_{12}^2=5$.

The $M=0$ lines for case (c) are plotted in Figs. 6–8 for $d_{13}^2/d_{12}^2=1, 3, 5$. It may be seen that growing wave solution exists, if the velocity of either of the waves 2 or 3 is near to that of wave 1. The maximum velocity difference between wave 1 and 2, 3 which still leads to complex roots increases as the coupling between 1 and 2, 3 increases. The asymptotes are:

$$\frac{p}{d_{12}} \to \infty \qquad \frac{r}{d_{12}} = \pm 2$$

$$\frac{r}{d_{12}} \to \infty \qquad \frac{p}{d_{12}} = 2\frac{d_{13}}{d_{12}}.$$

IV. Conditions of Complete Power Transfer

According to our boundary conditions, all the power is contained in wave 1 at $z=0$. In this section we shall investigate under what conditions this power can be completely transferred to waves 2 and 3.

Mathematically, it is equivalent to find the parameters which give $E_1(z)=0$. Since it does not appear that the general solution can be expressed in closed analytical form, we restrict generality and give only three solutions.

1) Complete power transfer is possible, if all three waves have the same velocity. Using the condition $\beta_1=\beta_2=\beta_3$, the amplitudes of the waves can be obtained from (6). Performing the calculations we get

$$E_1 = e^{-j\beta_1 z}\cos u_r z$$

$$E_2 = -jf_{12}d_{12}e^{-j\beta_1 z}\frac{\sin u_r z}{u_r}$$

$$E_3 = -jf_{13}d_{13}e^{-j\beta_1 z}\frac{\sin u_r z}{u_r} \tag{14}$$

where

$$u_r^2 = f_{12}d_{12}^2 + f_{13}d_{13}^2.$$

Thus complete power transfer takes place at the distance L, if

$$u_r^2 > 0 \text{ and } u_r L = (2k+1)\frac{\pi}{2}[k = 0, \pm 1, \pm 2 \cdots]. \tag{15}$$

2) Complete power transfer is possible, if the coupling coefficients are identical, the energy velocities of waves 2 and 3 are in the same direction, and the propagation coefficient of wave 1 is the arithmetical mean of the propagation coefficients of wave 2 and wave 3.

Using the conditions

$$\beta_1 = \frac{\beta_2+\beta_3}{2} \quad \text{and} \quad f_{12}d_{12} = f_{13}d_{13}, \tag{16}$$

the amplitudes can be written as follows:

$$E_1 = \frac{\exp(-j\beta_1 z)}{u_r^2}\{r^2 + 2f_{12}d_{12}^2\cos u_r z\}$$

$$E_2 = \exp(-j\beta_1 z)\frac{f_{12}d_{12}}{u_r^2}$$

$$\cdot\{r(1-\cos u_r z) + ju_r \sin u_r z\}$$

$$E_3 = \exp(-j\beta_1 z)\frac{f_{13}d_{13}}{u_r^2}$$

$$\cdot\{-r(1-\cos u_r z) + ju_r \sin u_r z\} \tag{17}$$

where

$$u_r^2 = r^2 + 2f_{12}d_{12}^2.$$

Complete power transfer takes place at the distance L, if

$$u_r^2 > 0, \qquad r^2 < 2d_{12}^2, \qquad \cos u_r L = -\frac{r^2}{2f_{12}d_{12}^2}. \tag{18}$$

3) Complete power transfer is still possible, when neither the velocities nor the coupling coefficients are identical, but then rather strict relationships apply between the quantities p, r, d_{12}, and d_{13}.

The mathematical conditions are as follows:

$$d_{12}^2 = \frac{1}{9f_{12}}\frac{p}{p-r}(2r-p)^2$$

$$d_{13}^2 = \frac{1}{9f_{13}}\frac{r}{r-p}(2p-r)^2. \tag{19}$$

Subject to the above conditions, the amplitudes of the waves can be written as follows:

$$E_1 = \tfrac{1}{2}\exp\left(-j\frac{\beta_1+\beta_2+\beta_3}{3}z\right)$$

$$\cdot\left\{1+\cos u_r z + j\tfrac{2}{3}(p+r)\frac{\sin u_r z}{u_r}\right\}$$

$$E_2 = \exp\left(-j\frac{\beta_1+\beta_2+\beta_3}{3}z\right)\frac{f_{12}d_{12}}{3u_r^2}$$

$$\cdot\{(2p-r)(1-\cos u_r z) - 3ju_r \sin u_r z\}$$

$$E_3 = \exp\left(-j\frac{\beta_1+\beta_2+\beta_3}{3}z\right)\frac{f_{13}d_{13}}{3u_r^2}$$

$$\cdot\{(2r-p)(1-\cos u_r z) - 3ju_r \sin u_r z\} \tag{20}$$

where

$$u_r^2 = \frac{2}{9}(2r-p)(r-2p).$$

Complete power transfer takes place at the distance L, if

$$u_r^2 > 0$$

and

$$u_r L = (2k+1)\pi[k = 0, \pm 1, \pm 2 \cdots]. \tag{21}$$

To see more clearly the relationship between the parameters, $f_{12}d_{12}^2/r^2$ and $f_{13}d_{13}^2/r^2$ are plotted in Fig. 9 against $a=p/r$. We can distinguish four regions.

1) $a<0, f_{12}=+1, f_{13}=+1$. All the waves in the same direction.
2) $0<a<\frac{1}{2}, f_{12}=-1, f_{13}=+1$. Wave 2 in the opposite direction.
3) $\frac{1}{2}<a<2$. No solution because $u_r^2<0$.
4) $a>2$, $f_{12}=1$, $f_{13}=-1$. Wave 3 in the opposite direction.

V. Examples

The Amplification Domain of a Traveling-Wave Tube

In the case of a traveling-wave tube, the circuit wave is coupled to both the slow and the fast space charge waves. Accordingly, we can identify wave 1 with the circuit wave, wave 2 with the fast wave, and wave 3 with the slow wave. Writing the propagation and coupling coefficients into the usual notations of traveling-wave theory[5,6] we obtain

$$\beta_1=\beta;\quad \beta_2=\beta_e(1-2C\sqrt{QC});\quad \beta_3=\beta_e(1+2C\sqrt{QC})$$

$$d_{12}=d_{13}=\frac{\beta_e C}{2\sqrt{QC}} \tag{22}$$

so that

$$\frac{p}{d_{12}}=-2b(QC)^{1/4}+4(QC)^{3/4}$$

$$\frac{r}{d_{12}}=-2b(QC)^{1/4}-4(QC)^{3/4}. \tag{23}$$

Since the coupling is the same to both space charge waves, the condition of amplification can be determined from Fig. 3. As p is always larger than r, the physically possible cases are below the $p/d_{12}=r/d_{12}$ line.

It may be seen from (23) that p/d_{12} and r/d_{12} are the functions of b and QC only. Therefore, the $b=$constant and $QC=$constant curves are plotted in Fig. 10, where for convenience the axes are rotated by 45 degrees. From the intersections with the $M=0$ lines the limiting values of b and QC can be determined. It may be seen that with decreasing values of QC the range of amplification is increased and pushed in the direction of lower values of b. The results of the two coupled waves theory[7] (taking account only of the circuit wave and the slow wave) are represented by the asymptotes (broken lines). If $QC>0.25$ the intersections with the asymptotes give good approximation.

The above results are, of course, familiar aspects of Pierce's theory of the traveling-wave tube;[5] the method of presenting the results here adopted is the "natural" one for the coupled wave picture, and shows up the essential physical phenomena from a different angle.

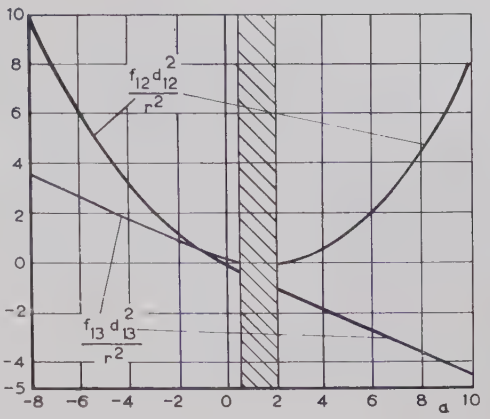


Fig. 9—A relationship between the parameters which results in complete power transfer.

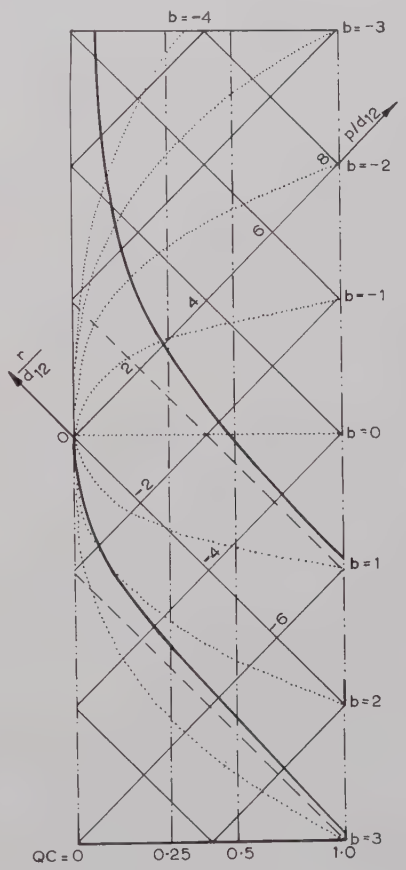


Fig. 10—The amplification domain of a traveling-wave tube. · · · $b=$constant curves. · · — · · — $QC=$constant curves.

[5] J. R. Pierce, "Travelling-Wave Tubes," D. Van Nostrand Co., Inc., New York, N. Y.; 1950.
[6] R. W. Gould, "Traveling-wave couplers for longitudinal beam-type amplifiers," Proc. IRE, vol. 47, pp. 419–429; March, 1959.
[7] R. W. Gould, "A coupled mode description of the backward wave oscillator and the Kompfner dip condition," IRE Trans. on Electron Devices, vol. ED-2, pp. 37–42; October, 1955.

Traveling-Wave Tube at Kompfner Dip

For certain values of beam voltage and current, the power of the circuit wave can be completely transferred to the space charge waves.[8] This is known as the Kompfner dip condition and it is useful because it permits a direct measurement of the traveling-wave tube parameters. A number of authors[6,9,10] have published numerical solutions for the location of the point.

Unfortunately, our formulas derived in Section IV are not generally applicable owing to the severe restrictions represented by (19). It turns out that its validity is restricted to one particular case, namely to

$$b = -\frac{3}{2} \quad \text{and} \quad QC = \frac{5}{16}.$$

The amplitudes of the three waves then can be obtained from (20). Since this is the only analytical solution found so far for the Kompfner dip condition, it seems to be worthwhile to write up the formulas.

$$|E_1|^2 = \tfrac{1}{4}\{(1 + \cos \sqrt[4]{20}\, d_{12}z)^2 + 2(\sin \sqrt[4]{20}\, d_{12}z)^2\}$$

$$|E_2|^2 = \frac{3\sqrt{5}+5}{40}(1 - \cos \sqrt[4]{20}\, d_{12}z)^2 + \frac{1}{\sqrt{20}}(\sin \sqrt[4]{20}\, d_{12}z)^2$$

$$|E_3|^2 = \frac{3\sqrt{5}-5}{40}(1 - \cos \sqrt[4]{20}\, d_{12}z)^2 + \frac{1}{\sqrt{20}}(\sin \sqrt[4]{20}\, d_{12}z)^2. \tag{24}$$

Complete power transfer takes place when

$$\sqrt[4]{20}\, d_{12}z = (2k+1)\pi \quad [k = 0, \pm 1, \pm 2 \cdots].$$

A feature of this solution (unlike others encountered) is that the variation of the amplitudes with distance is periodic.

Complete Power Transfer in Waveguides

Let us consider three coupled waveguides, where all the phase and energy velocities are in the same direction and waveguides 2 and 3 are not coupled.

A practical example might take the form of a power divider in which the power in 1 is transferred to particular modes in 2 and 3 in a predetermined ratio. In this example we assume that

$$\frac{\beta_2}{\beta_3} \quad \text{and} \quad \frac{|E_2(L)|^2}{|E_3(L)|^2} = s \tag{25}$$

are given, and β_1/β_3 and d_{13}/d_{12} are to be found.

It can be shown from (19), (20), and (25) that

$$s = -a. \tag{26}$$

Having obtained the value of a, the ratio d_{12}/d_{13} can be calculated from (19), or from Fig. 9. From the definitions of p and r [Eq. (11)] we get furthermore

$$\frac{\beta_1}{\beta_3} = \frac{1}{1-a}\left[1 - a\frac{\beta_2}{\beta_3}\right]. \tag{27}$$

Assuming for example, $\beta_2/\beta_3 = 0.8$ and $s = 0.5$ we get $\beta_1/\beta_3 = 0.93$ and $d_{12}/d_{13} = 0.89$. The power in the waveguides for the above values of the parameters is shown in Fig. 11 as a function of the normalized distance $d_{12}z$.

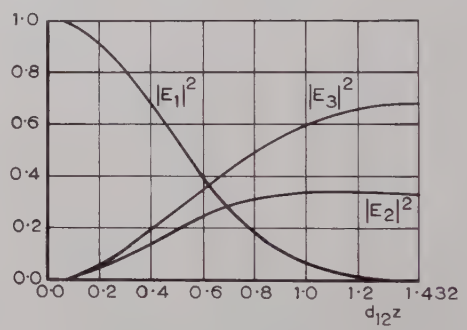


Fig. 11—The power in the waveguides as a function of the normalized distance, $d_{12}z$.

Double Beam Backward-Wave Oscillator

In a backward-wave oscillator utilizing two separate electron beams[11] so disposed that the interaction between them may be neglected, five waves are playing essential roles: the backward circuit wave, the two fast, and the two slow waves. However, if QC is large enough the problem can be greatly simplified. It is sufficient then to take account of the interaction of the backward circuit wave with the two slow waves.

Thus we can identify wave 1 with the backward circuit wave, and waves 2 and 3 with the slow waves. Since all the energy velocities are in the same direction, $f_{12} = f_{13} = 1$. The tube will oscillate if the power contained in the slow waves can be completely transferred to the backward circuit wave.

Let us investigate first the simplest case, when both beams (and thus both slow waves) are identical. Complete power transfer takes place if the propagation coefficient of the backward circuit wave is equal to those

[8] R. Kompfner, "On the operation of the traveling wave tube at low level," *J. Brit. IRE*, vol. 10, pp. 283–289; August–September, 1950.

[9] H. R. Johnson, "Kompfner dip conditions," PROC. IRE, vol. 43, p. 874; July, 1955.

[10] R. D. Weglein, "Backward wave oscillator starting conditions," IRE TRANS. ON ELECTRON DEVICES, vol. ED-4, pp. 177–179; April, 1957.

[11] E. A. Ash and A. C. Studd, "Multiple Beam Backward Wave Oscillators," presented at the Electron Tube Conf., Mexico City, Mexico; June, 1959.

of the slow waves, and the relation $(d_{12})_2 L = \pi/2\sqrt{2}$ applies (15), where $(d_{12})_2$ is the coupling coefficient between the backward circuit wave and one of the slow waves at the start of oscillation. If there is only a single beam, the condition of start oscillation[7] is $(d_{12})_1 L = \pi/2$, where $(d_{12})_1$ is the coupling coefficient between the backward circuit wave and the slow wave. Thus the necessary value of the coupling coefficient for the start of oscillation is smaller if both beams are present. Keeping the beam voltage constant, the ratio of the starting currents is as follows:

$$\frac{I_2}{I_1} = \left[\frac{(d_{12})_2}{(d_{12})_1}\right]^4 = \frac{1}{4}. \tag{28}$$

Thus the beam current in a double-beam backward-wave oscillator drops by a factor 4, and the total current is still only half of that which is necessary in the single beam device.

If the beam voltages are slightly different, the coupling coefficients can still be regarded as identical because they are slowly varying functions of the beam voltage. Thus, applying the formulas of Section IV, the propagation coefficient of the backward circuit wave is the arithmetical mean of the propagation coefficients of the slow waves.

It may be seen from (17) that for finite voltage differences the starting current increases, which agrees qualitatively with the experimental results.[11] If

$$r^2 > 2d_{12}^2,$$

the amplitude of the backward circuit wave cannot be made zero. Thus, beyond a certain voltage difference, no oscillation can be obtained, however long the circuit.

ACKNOWLEDGMENT

The author wishes to thank Dr. E. A. Ash for many interesting discussions. Thanks are also due to Standard Telecommunication Laboratories Ltd., for permission to publish the paper.

Noise Figures of Reflex Klystron Amplifiers*

KORYU ISHII†

Summary—**The noise figure of the 2K25 reflex klystron amplifier was investigated. The noise figure of the reflex klystron amplifier depends on operating frequency, electronic impedance, circuit impedance, and operating electronic mode. Experimental results show that a noise figure of 5 db is possible under particularly carefully adjusted conditions. In order to obtain the low-noise figure, careful electronic tuning and the impedance adjustments are particularly important. Generally, relatively low noise figures were obtained when the electronic tuning was good. Noise figures of cascaded reflex klystron amplifiers were also investigated experimentally. Noise figures of the cascaded amplifier were generally higher than that of the single stage amplifier, but still low enough to use this reflex klystron amplifiers as a preamplifier of a microwave receiver to increase the sensitivity of the receiving system.**

INTRODUCTION

THE use, as regenerative or negative conductance amplifiers, of reflex klystrons originally designed for use in oscillators, would offer several advantages to microwave receiver design. Ordinary, small-power reflex klystrons are relatively inexpensive, and require neither the high voltages used in TW tubes nor the great magnetic force necessary in magnetrons.

There is some controversy about such an application for reflex klystrons. In the first place, it is questioned whether employment of the reflex klystron amplifier really does increase the sensitivity of a microwave receiver. To increase the receiver's sensitivity, the reflex klystron would have to provide a good gain and at the same time have a low noise figure.

Several papers have been published describing the gain achieved with reflex klystron amplifiers. Okabe[1] obtained a gain of over 20 db at 3000 mc with a 707B reflex klystron. Ishii[2,3] obtained a gain of more than 16 db at 9760 mc with a 723A/B reflex klystron. Quate, Kompfner and Chisholm[4] reported a gain of more than 30 db at 11,000 mc with a WE445A reflex klystron. These papers demonstrate that a substantial gain improvement is possible, but no useful data on noise figures was obtained. For example, Okabe reported a noise figure of less than 7 db but Quate reported 40 db. Clearly, a study of the noise figure itself was required if the value of the reflex klystron amplifier was to be verified or denied.

* Manuscript received by the PGMTT, September 4, 1959; revised manuscript received, November 9, 1959.

† Dept. of Elec. Engrg., Marquette University, Milwaukee, Wis.

[1] T. Okabe, "Microwave amplification by the use of reflex klystron," *Report of Microwave Research Committee in Japan;* June and July, 1952.

[2] K. Ishii, "*X*-band receiving amplifier," *Electronics*, vol. 28, pp. 202–210; April, 1955.

[3] K. Ishii, "Oneway circuit by the use of a hybrid *T* for the reflex klystron amplifier," PROC. IRE, vol. 45, p. 687; May, 1957.

[4] C. F. Quate, R. Kompfner, and D. A. Chisholm, "The reflex klystron as a negative resistance type amplifier," IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 173–170; July, 1958.

The noise level of the reflex klystron amplifier is dependent on the conditions of the tube and circuit. In order to study one reflex klystron amplifier under various conditions, the circuit shown in Fig. 1 was used. This amplifier used a 2K25 (723A/B) and operated in the X band. The reflex klystron, as shown in the figure, is mounted in the middle of the waveguide, and the output impedance to the reflex klystron is adjusted with six screw tuners and one coaxial shorting plunger to obtain suitable regenerative feedback. Input signals are fed into the left opening of the waveguide and output power is taken from the right end.

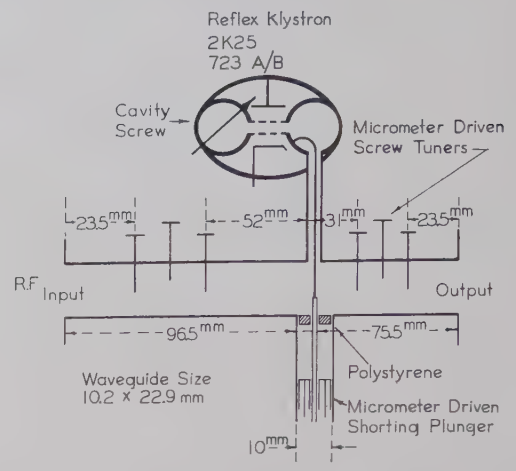


Fig. 1—Schematic diagram of the 2K25 reflex klystron amplifier circuit.

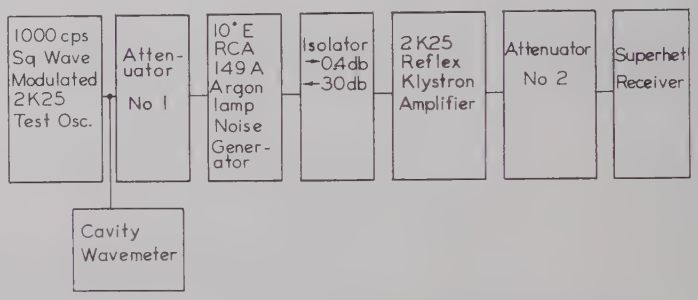


Fig. 2—Noise measurement setup for reflex klystron amplifier.

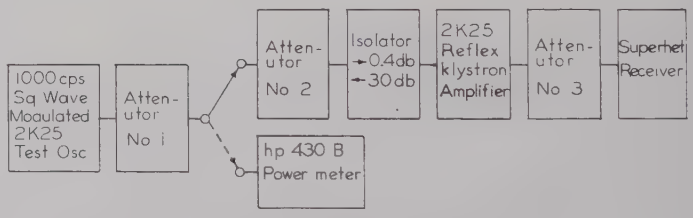


Fig. 3—Small signal method.

Noise Figure Measurement

Two methods were used to measure noise, one of which was primarily for checking results. The first was the noise generator method, as shown in Fig. 2. The alternate procedure, used for vertification of the other, was the small signal method, shown in Fig. 3.

The regenerative properties of the reflex klystron make it necessary that the same conditions of amplification be maintained for all noise measurements. The external circuits and tube supply voltages have a great effect on the noise figure of the reflex klystron amplifier.

Therefore, extreme precautions were taken toward supply voltages and the external circuit impedance to keep the same gain of the amplifier during the experiment. Power supply to the reflex klystron anode was electronically stabilized and a battery was used to give stable voltage to the repeller of the reflex klystron. Signal source impedance was carefully examined before the experiment by a standing wave detector. The VSWR was approximately 1.15. In the noise generator method of Fig. 2, due to the isolation effect of the isolator, no detectable change in VSWR on the input waveguide of the amplifier was observed by noise lamp on and off conditions. Similar impedance stabilizing effect of the isolator was observed in the small signal method of Fig. 3. As a matter of fact, the amplifier, carefully adjusted under these conditions, was very stable, and showed very good reproducibility of the same gain after various kinds of circuit manipulations such as noise lamp off and on, or replacing to a reflectionless termination. The amplifier showed very good linearity for small signals. Whenever the equal external impedance was given, the amplifier reproduced the same gain.

Using the noise generator method, an oscillator signal was fed into the amplifier through the attenuator, the argon discharge noise generator, and an isolator. Simultaneously, a noise signal was amplified in the same manner. Then the amplifier was adjusted for optimum gain and the gain was measured with the attenuator. The oscillator was shut off, leaving only the noise signal under amplification. At this time, the amplifier output was read with a thermocouple meter connected to the output of the IF amplifier of the RF head of a USAF APS-3 radar receiver. Though it was not shown in Figs. 2 and 3, the receiver contained an isolator at the input in this case to prevent the local oscillator disturbance to the 2K25 amplifier. This isolator also had an effect to avoid input temperature problem of the receiver. Maintaining the same conditions as above, the noise generator which was isolated by the isolator was replaced by a reflectionless termination and measurements were again taken. The VSWR measurement before the experiment showed there was no detectable change in VSWR after replacing the noise generator by the reflectionless termination. The over-all noise figure of the amplifier and receiver is given by

$$F_{12} = 15 + 10 \log \frac{M}{N - M} \text{ (db)} \qquad (1)$$

where 15 db is the excess noise of the RCA 149 argon lamp with forward isolator insertion loss subtracted, N is the reading of the thermocouple meter when the noise generator was connected, and M is the reading with a

reflectionless termination. The isolator helps hold output impedance constant for the reflex klystron.

The noise figure of the reflex klystron amplifier, then, is

$$F_1 = F_{12} - \frac{F_2 - 1}{G_1} \qquad (2)$$

where F_2 is the noise figure of the receiver alone, and G_1 is the gain of the reflex klystron amplifier. In this case, F_2 is measured under the same conditions as F_{12}. To properly assess the noise figure of the klystron amplifier, the contribution of load noise should be subtracted from the experimental value. The load noise contribution was, however, minor because of the isolators, one of which was in front of the amplifier and another in front of the receiver.

The small signal method was used to verify results of the noise generator experiments. The schematic diagram for this method is shown in Fig. 3. The isolator in front of the amplifier, as before, keeps the output impedance of the reflex klystron constant during measurement. The noise figure is given by the equation

$$F_{12} = P_s(\text{dbw}) - \{-204(\text{dbw}) + B_{12}(\text{db})\} \qquad (3)$$

where P_s is the available signal input power to the reflex klystron amplifier required in order to double the noise power output, and B_{12} is the over-all noise bandwidth.

It is important to measure P_s carefully, and if residual reflections exist, correction must be made. Image effect is not serious in this case because the reflex klystron bandwidth is very narrow in comparison with the receiver IF frequency. Difficulty in obtaining accurate measure of P_s made this small signal method difficult and, for this reason, the noise generator method was used as the primary means of investigation.

Noise Figure of Reflex Klystron Amplifier

The measurements made of the 2K25, as described above, reveal that circuit impedances and supply voltages for the reflex klystron affect the noise figure significantly. For example, if the operating frequency is changed, output impedance must be adjusted and electronic impedance must follow to maintain optimum gain; as a result, the noise figure changes. Table I shows the relation between the noise figure and the operating frequency. For each frequency, amplifier circuit conditions were adjusted for optimum gain by means of the screw tuners, cavity screw, shorting plunger, and adjustments of anode and repeller voltages. The circuit adjustments were not difficult to make and noise figures ranged from 5.2 db to 23.2 db, depending on operating conditions.

Table II shows noise figures for a fixed frequency level, with the reflex klystron electron transit cycles in a number of operating modes. Noise figures varied from 15.5 db to 24.8 db.

TABLE I

Noise Figure of 2K25 Amplifier under Various Operating Frequencies

Operating Frequency mc	Frequency Bandwidth mc	Noise Figure db	Optimum Gain db
9310	7.6	15.5	22
9351	6	10.5	21
9362	6.5	5.2	16
9387	11.5	7.2	15
9407	6.2	6.2	17
9429	14	9	20
9452	12	23.2	16
9499	5.2	15	15

TABLE II

Noise Figures of 723A/B Amplifier under Various Operating Modes at 9310 mc

Mode, Electron Cycles N	Optimum Gain db	Frequency Bandwidth mc	Noise Figure db
$8\frac{3}{4}$	18	5.4	17.8
$9\frac{3}{4}$	25	22	24.8
$10\frac{3}{4}$	17	34	17.5
$11\frac{3}{4}$	22	7.6	15.5

TABLE III

Noise Figures of 723A/B Amplifier under Various Circuit Conditions Operated in $N=8\frac{3}{4}$ Mode at 9429 mc

Case	Optimum Gain db	Frequency Bandwidth mc	Noise Figure db
A	13	30	12.5
B	22	19.6	11
C	29	10	20

Table III lists noise figures measured with the reflex klystron amplifier operated at one frequency and in one mode, but with varying circuit conditions. Again, noise figures varied. For this experiment, the 2K25 was operated in mode $N=8\frac{3}{4}$ and at frequency 9429 mc. With varying circuit conditions, repeller and anode voltages were adjusted for optimum gain. Cases A, B, and C in the table each represent a different adjustment of the screw tuner, cavity screw, and shorting plunger.

Finally, when the reflex klystron was operated in one mode and at one frequency, circuit conditions were unchanged. However, the repeller voltage, or electronic tuning, was changed, and the noise figure varied, as shown in Fig. 4. With repeller voltage higher than -152 volts, the electronic tuning is relatively poor, gain is low, and noise is high. But with repeller voltage at -153 volts, electronic tuning is fair and the noise figure is minimum. When repeller voltage goes below -154 volts, electronic impedance approaches the self-oscillation region and gain increases, but the noise figure rises, too. Both the noise generator method and the small signal method were used for noise measurements and both produced similar results.

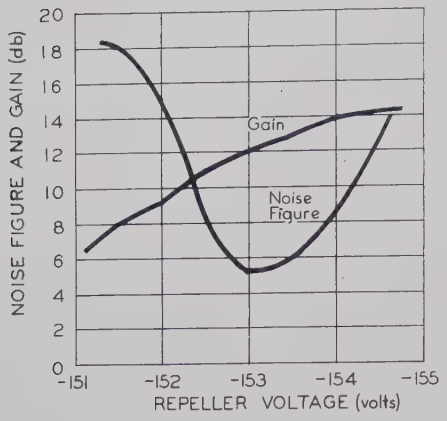


Fig. 4—An example of noise figure, gain and repeller voltage relations of 723A/B amplifier.

If careful adjustment of circuit and electronic impedances can reduce noise figures to 5 db, the use of a reflex klystron amplifier as a preamplifier will increase the sensitivity of a microwave receiver. In this experiment the RF head of a USAF APS-3 aircraft radar receiver was used. It has a sensitivity of −98 dbm and a noise figure of 28 db. A 2K25 amplifier with a 14-db gain and a 5-db noise figure was connected in front of the receiver. The resulting amplifier-receiver could "see" a signal of 9362 mc, 10 db lower than could the APS-3 receiver alone. The sensitivity, then, was improved by −98 dbm to −108 dbm.

Assuming that a reflex klystron amplifier will increase sensitivity, cascading two of them should improve receiver operation even more. In Table IV, results of experiments with a cascaded reflex klystron amplifier are given. In this table, *I*, *P*, and *T* represent methods of combining two reflex klystron amplifiers by means of an isolator (*I*), a phaseshifter (*P*), and a transformer (*T*). *P-T* and *I-P-I* represent combinations of these elements. The APS-3 receiver was taken as a reference-sensitivity of a receiving system. The results of the experiment show that receiver sensitivity is increased, and that the noise figure can be reduced enough to make the reflex klystron amplifier acceptable as a preamplifier for ordinary microwave receivers. Generally, over-all noise of the cascaded amplifier is higher than that of a single stage one, but the additional gain is high enough to give a net improvement in sensitivity.

When the 2K25 amplifier, which has an isolator ahead of it, if gain G_1, and noise figure F_1 are connected in front of a receiver of noise figure F_2 by an isolator, the over-all noise figure is given by

$$F_{12} = F_1 + \frac{F_2 - 1}{G_1} . \qquad (4)$$

For example, when the isolator-phase shifter-isolator coupled amplifier is connected in front of the APS-3 receiver which contained an isolator in its input, then

$$F_1 = 16 \text{ db}, \quad F_2 = 28 \text{ db}, \quad \text{and} \quad G_1 = 60 \text{ db};$$

TABLE IV

Sensitivities Improved by the Use of Reflex Klystron Amplifiers

Reflex Klystron Amplifier	Frequency Bandwidth mc	Noise Figure db	Gain db	Sensitivity Minimum Detectable Signal Below APS-3 Receiver db
APS-3 Receiver	2	28	0	0
	(Reference Standard)			
Single Stage	20	5	14	10
	(Following are two stage amplifiers)			
Direct Coupled	6	28	30	4
Isolator Coupled	4.4	26	28	4
Phase Shifter Coupled	2	16	43	19
I-P Coupled	4.5	8	35	26
P-I Coupled	2	17.5	42	26
P-I-I Coupled	3.25	11	49	26
I-I-P	3.7	13	54	28
I-I	2	5	31	12
I-P-I	2.4	16	60	42
I-I-I	1.6	17	49	16

and

$$F_{12} \doteq F_1 = 16 \text{ db}.$$

Thus, the over-all noise figure is improved from 28 db to 16 db.

Conclusions

Since the noise figure of the reflex klystron is dependent upon the operating output impedance, the frequency, and the electronic impedance, it is difficult to predict what its value will be. However, measurements of the 2K25 amplifier indicates that a noise figure on the order of 6.5 db can be obtained for an optimum net gain of 24.5 db at 9362 mc. And if the optimum gain required is below 15 db, the noise figure can be decreased to approximately 5 db if circuit conditions are carefully adjusted.

Minimum noise figures are usually obtained when the circuit is adjusted for optimum gain or when good electronic tuning was performed in individual electronic modes. If electronic tuning is poor, or the tube is operated in semioscillation conditions, the noise figure is high.

The 2K25 reflex klystron amplifier has a low enough noise figure to be used as a preamplifier to improve the sensitivity of ordinary microwave receivers.

Acknowledgment

The author wishes to thank Prof. E. H. Scheibe of the University of Wisconsin for his friendly advice in this research and S. Krupnik of Marquette University, and P. Smith of A.C. Spark Plug Co., for help in the preparation of this manuscript.

On Measurements of Microwave $\overline{E}$ and $\overline{H}$ Field Distributions by Using Modulated Scattering Methods*

MING-KUEI HU†

Summary—The modulated scattering method of Justice, Rumsey, and Richmond for measuring $\overline{E}$ field distribution is extended to the measurement of $\overline{H}$ field distribution by using a loop scatterer formed by two diodes. This diode loop method has the particular advantage of eliminating the large and undesirable effect produced by the associated $\overline{E}$ field when measuring the $\overline{H}$ field.

A scattering analysis of the modulated diode loop is presented. It explains the principle of this new method and also supports the advantage mentioned above. A similar analysis for the modulated diode scatterer used in measuring $\overline{E}$ is also presented. It is believed that the explanation based upon this analysis for the $\overline{E}$ measurement is more satisfactory than that given by Richmond which is based upon a qualitative description of the diode scatterer.

A THEORETICAL basis and some experimental results on an ingenious scattering method for measuring $\overline{E}$ field were published by Justice and Rumsey.[1] The method was later modified by Richmond[2] with the use of a modulated diode as the scatterer. This modulated scattering method has been extended by the present author to the measurement of $\overline{H}$ field by using a loop scatterer formed by two diodes. This diode loop method has the particular advantage of eliminating the large and undesirable effect produced by the associated $\overline{E}$ field when measuring the $\overline{H}$ field. Such $\overline{E}$ field effect is really what makes most of the $\overline{H}$ field measurements unreliable.

A scattering analysis of the modulated diode loop has been worked out. It explains the working principle of this new method for measuring $\overline{H}$ and also supports the advantage mentioned above. A similar analysis has also been applied to the modulated diode scatterer used by Richmond. An explanation, based upon this analysis, for the working principle of the modulated scattering method for measuring $\overline{E}$ is believed to be more satisfactory and enlightening. For the convenience of presentation, the simpler analysis of a modulated diode is given in Section I. The analysis of a modulated diode loop and its application to the measurement of $\overline{H}$ field are presented in Section II.

* Manuscript received by the PGMTT, August 20, 1959; revised manuscript received, November 12, 1959. This study was supported by Rome Air Dev. Center, Rome, N. Y., under Contract AF 30(602)-1640.

† Dept. of Elec. Engrg., Syracuse University, Syracuse, N. Y.

[1] R. Justice and V. H. Rumsey, "Measurement of electric field distributions," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-3, pp. 177–180; October, 1955.

[2] J. H. Richmond, "Measurement of field distributions," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-3, pp. 13–15; July, 1955.

I. Back-Scattered Signal From a Diode Dipole and $\overline{E}$ Measurement

The method of analysis is an extension of that used by Y. Y. Hu in a paper on back-scattering cross section.[3] In Fig. 1, with the effect of the slightly conducting leads for modulation neglected, the diode with its two short conducting leads is considered as a short dipole center loaded with the diode junction impedance. The two terminals across the source which produces the field to be measured and the two terminals across the diode junction are considered as a two-port configuration. The two pairs of terminals of this two-port will be referred to as the source terminals and the dipole terminals in the following analysis.

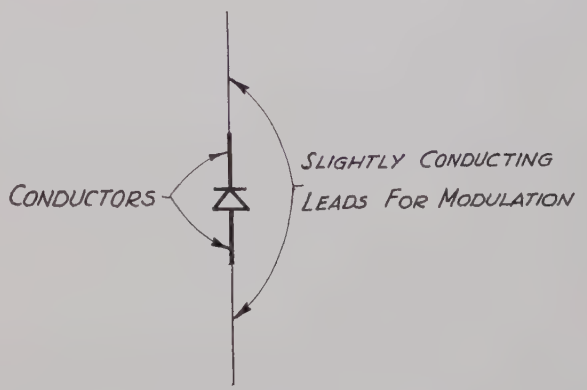


Fig. 1—A diode dipole scatterer.

It is clear that, with the diode junction impedance removed, the remaining part is linear; therefore, the principle of superposition applies. As far as the junction impedance is concerned, it serves only as a link between the dipole terminal current and the dipole terminal voltage, and the modulation merely changes the value of this junction impedance. Based upon such an approach, the back-scattered signal at the source terminals due to a loaded dipole can always be expressed in terms of the solutions of the following two simpler but more basic problems.

1) The back-scattered signal of an open-circuited dipole: If the dipole is located not too close to the source

[3] Y. Y. Hu, "Back-scattering cross section of a center-loaded cylindrical antenna," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 140–148; January, 1958.

structure, then the higher order effects between the dipole and the source can be neglected. Let $\bar{E}(\bar{r})$ be the field to be measured, and $\bar{I}_0(\bar{r})$ be the current induced on the dipole by $\bar{E}(\bar{r})$, then the electric field intensity produced by $\bar{I}_0(\bar{r})$ can be expressed by the following line integral:

$$\int_{\text{scatterer}} \bar{I}_0(\bar{r}') \cdot \bar{\bar{G}}(\bar{r}', \bar{r}) ds'.$$

where $\bar{r}$ and $\bar{r}'$ are the position vectors of the field point and the source point respectively. $\bar{\bar{G}}(\bar{r}', \bar{r})$ is the dyadic Green's function[4] of the dipole with the effect of the source structure neglected. The differential length element ds' indicates that the line integral is to be evaluated with respect to the source coordinates. If s (also s') is considered as a length variable measured along the dipole with $s=0$ at the dipole terminals (these terminals are assumed to be located very close to each other), then there clearly exists a one-to-one correspondence between s and $\bar{r}$. Therefore, $\bar{E}(\bar{r})$ may also be written as $\bar{E}(s)$ for values on the dipole, and the above line integral as

$$\int_{\text{scatterer}} \bar{I}_0(s') \cdot \bar{\bar{G}}(s', s) ds'.$$

These expressions in terms of the variables s and s' should be considered as obtained from a transformation of variables from $\bar{r}$ to s, rather than as a direct substitution of s for $\bar{r}$. Using the boundary condition that the tangential component of the total electric field intensity is zero on the dipole, we then have the following relation over the dipole except at the feed:

$$\bar{E}(s) \cdot \hat{s} + \int_{\text{scatterer}} \bar{I}_0(s') \cdot \bar{\bar{G}}(s', s) \cdot \hat{s} ds' = 0$$

where $\hat{s}$ is a unit tangent vector along the dipole. Using $\hat{s}$, $\bar{I}_0(s)$ may also be written as $I_0(s)\hat{s}$ with $I_0(s)$ as the scalar part of $\bar{I}_0(s)$. At the feed, the value of the above line integral becomes infinity. But if we integrate the left-hand side of the above relation with respect to s over a very small interval from $0-\epsilon$ to $0+\epsilon$, then the first term gives zero and the second term gives a finite value. This finite valued quantity, in fact, is the voltage, V_{10}, induced across the dipole terminals by the electric field intensity to be measured. The above relations satisfy the conditions required for the definition of the Dirac delta function $\delta(s)$; therefore we have

$$V_{10}\delta(s) = \bar{E}(s) \cdot \hat{s} + \int_{\text{scatterer}} \bar{I}_0(s') \cdot \bar{\bar{G}}(s', s) \cdot \hat{s} ds'. \quad (1)$$

It is obvious that the open circuit condition implies

$$I_0(0) = 0. \quad (2)$$

[4] In this dipole case, a scalar Green's function may be used.

For simplicity in notation, a simplified form of integral expression is defined as follows:

$$\int \bar{I}_0 \cdot \bar{\bar{G}} \cdot \hat{s} = \int_{\text{scatterer}} \bar{I}_0(s') \cdot \bar{\bar{G}}(s', s) \cdot \hat{s} ds'.$$

Similar notations will be used throughout this paper.

Using the reciprocity theorem, the back-scattered signal V_0 at the source terminals can be expressed in terms of $\bar{I}_0(s)$ as,

$$V_0 = \frac{1}{I} \left\{ \int \bar{E} \cdot \bar{I}_0 \right\}. \quad (3)$$

2) The signal received from a radiating dipole: If a current I_1 is fed into the dipole terminals, the dipole will radiate. Using the same assumption as given in 1), the following relation is obtained along the dipole:

$$V_{11}\delta(s) = \int \bar{I}_1 \cdot \bar{\bar{G}} \cdot \hat{s} \quad (4)$$

where $\bar{I}_1(s) = I_1(s)\hat{s}$ is the radiating current distribution on the dipole, and V_{11} is the voltage across the dipole terminals due to the feeding current I_1. It is obvious we also have

$$I_1(0) = I_1. \quad (5)$$

A simple relation which is useful later can be obtained by multiplying (4) by $I_0(s)$ and integrating over the dipole scatterer,

$$\int\int \bar{I}_1 \cdot \bar{\bar{G}} \cdot \bar{I}_0 = 0. \quad (6)$$

The relation can be justified with the use of (2).

Similarly as in 1), the signal V_r received at the source terminals is given by

$$V_r = \frac{1}{I} \left\{ \int \bar{E} \cdot \bar{I}_1 \right\}. \quad (7)$$

Now in the general case of back-scattered signal due to a loaded dipole, both $\bar{I}_0(s)$ and $\bar{I}_1(s)$ exist simultaneously on the dipole; therefore, the signal V received is the sum of V_0 and V_r.

$$V = V_0 + \frac{I_1}{I} \int \bar{E} \cdot \tilde{I}_1 \quad (8)$$

where $\tilde{I}_1(s)$ is defined by

$$\tilde{I}_1(s) = \frac{\bar{I}_1(s)}{I_1(0)} = \frac{\bar{I}_1(s)}{I_1}. \quad (9)$$

Evidently $\tilde{I}_1(0)$ has unit magnitude.

At the same time, in this loaded dipole case the dipole terminal voltage V_1 is related to the dipole terminal current I_1 by the diode junction impedance Z_L as follows:

$$V_1 = -Z_L I_1. \quad (10)$$

On the other hand, by using the principle of superposition, we have

$$V_1 = V_{10} + V_{11}. \tag{11}$$

Combining (1), (4), (10), and (11), the following relation is obtained:

$$-Z_L I_1 \delta(s) = \overline{E}\cdot\hat{s} + \int \bar{I}_0 \cdot \overline{\overline{G}} \cdot \hat{s} + \int \bar{I}_1 \cdot \overline{\overline{G}} \cdot \hat{s}. \tag{12}$$

This equation describes the boundary condition along the dipole in the loaded case. Multiplying (12) by $I_1(s)$ and integrating over the dipole, we have

$$-Z_L I_1 = \int \overline{E}\cdot\check{I}_1 + I_1 \int\int \check{I}_1 \cdot \overline{\overline{G}} \cdot \check{I}_1. \tag{13}$$

In deriving the above equation, the relations (6) and $\bar{I}_1(s) = I_1 \check{I}_1(s)$ are used. The double integral in the above equation is known as the input impedance, Z, of the dipole;[3] therefore we have

$$I_1 = -\frac{\int \overline{E}\cdot\check{I}_1}{Z + Z_L}. \tag{14}$$

Substituting (14) into (8), V is now given by

$$V = V_0 - \left(\frac{1}{I}\right)\frac{\left\{\int \overline{E}\cdot\check{I}_1\right\}^2}{Z + Z_L}. \tag{15}$$

When (15) is applied to the modulated scattering method of measuring $\overline{E}$, it is reasonable to assume that both $\overline{E}$ and $\hat{s}$ are constant along the diode dipole. This assumption is justified by the fact that in measurement the dipole is chosen to be small in comparison with the wavelength. Therefore the factor $\overline{E}\cdot\hat{s}$ can be taken outside of the integral $\int \overline{E}\cdot\check{I}_1$, and the remaining factor can be recognized as the effective length l_d of the dipole. $\overline{E}\cdot\hat{s}$ is clearly the component of $\overline{E}$ along the direction of the diode, if it is denoted by E_d, and we then have

$$\int \overline{E}\cdot\check{I}_1 = E_d l_d. \tag{16}$$

Finally, the back-scattered signal, V, from a diode dipole is obtained from (15) as

$$V = V_0 - K\frac{E_d^2 l_d^2}{Z + Z_L} \tag{17}$$

with $K = 1/I$. In this equation, the first term is a continuous signal; therefore, it is not detected by the coherent detection system described by Richmond. If the source excitation is kept fixed, K is a constant. The only quantity which is affected by the modulation is the diode junction impedance Z_L appearing in the second term. Therefore the signal due to the second term is modulated and is detected. This shows clearly that the $\overline{E}$ field can be measured by using such a modulated diode as a scatterer.

It should be noted that the modulation used is generally of the on-and-off type. As a result, the diode junction impedance Z_L has essentially two different values—one for the off-period and one for the on-period. From (17), it can be seen that even during the off-period the back-scattered signal due to the second term is not zero, but zero value was used in Richmond's analysis of the coherent detection system. However, with slight modification in his analysis, the following two properties can still be shown:

a) The magnitude of the signal detected in the coherent system is still proportional to the square of the magnitude of the $\overline{E}$ component along the direction of the diode.

b) The phase of the detected signal is now equal to twice the phase angle of $\overline{E}$ plus a constant phase shift. This phase shift depends upon Z and the two different values of Z_L.

II. Back-Scattered Signal From a Diode Loop and $\overline{H}$ Measurement

The scattering loop is formed of two diodes connected as shown in Fig. 2. The loop is assumed to be symmetrical about a line passing through the two diode junctions and also symmetrical about a line passing through the two midpoints on the loop between the two diode junctions. The modulation is applied through a pair of slightly conducting leads to the two diodes in parallel; it modulates both diode junction impedances simultaneously. The applied modulation merely changes the values of these two impedances. As far as the microwave is concerned, the diode loop can be considered simply as an ordinary conducting loop loaded with two impedances due to the two diode junctions.

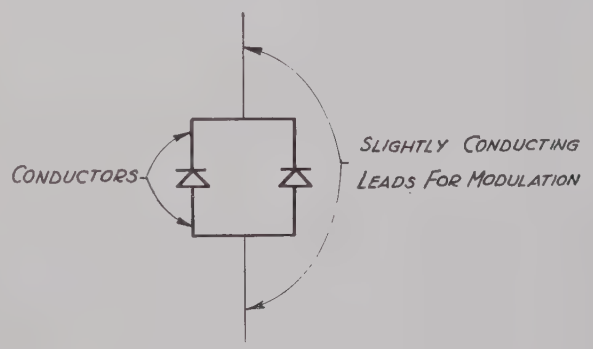


Fig. 2—A diode loop scatterer.

The analysis of a loaded loop is a further extension of that for the loaded dipole, but it is more involved. For simplicity of presentation, some of the obvious assumptions will still be used but not stated explicitly in this analysis. The two terminals of the source and the two pairs of terminals across the two diode junctions are

considered as a three-port configuration. The three pairs of terminals of this three-port will be referred to as the source terminals, the loop terminals 1-1, and the loop terminals 2-2. Similarly, as in the case of a loaded dipole, the back-scattered signal at the source terminals due to a loaded loop can be expressed in terms of the following three problems:

1) The back scattered signal of an open-circuited loop. The following equation which corresponds to (1) is obtained:

$$V_{10}\delta(s - s_1) + V_{20}\delta(s - s_2) = \bar{E}\cdot\hat{s} + \int \bar{I}_0\cdot\bar{\bar{G}}\cdot\hat{s} \quad (18)$$

where $\bar{\bar{G}}$ is now the dyadic Green's function for the loop. s is a length variable measured along the loop, but with $s=s_1$ and $s=s_2$ denoting the positions of the two pairs of loop terminals 1-1 and 2-2 respectively. V_{10} and V_{20} are the voltages induced across the loop terminals 1-1 and 2-2. Instead of (2) we now have the following two equations:

$$\begin{aligned} \bar{I}_0(s_1) &= 0 \\ \bar{I}_0(s_2) &= 0. \end{aligned} \quad (19)$$

Eq. (3) still holds, and is repeated here for convenience

$$V_0 = \frac{1}{I}\left\{\int \bar{E}\cdot\bar{I}_0\right\}. \quad (20)$$

2) The signal received from a radiating loop excited by a current I_1 at terminals 1-1 but with terminals 2-2 open. A modified form of (4), for this case, is the following:

$$V_{11}\delta(s - s_1) + V_{21}\delta(s - s_2) = \int \bar{I}_1\cdot\bar{\bar{G}}\cdot\hat{s} \quad (21)$$

where V_{11} and V_{21} are the voltages across terminals 1-1 and 2-2 respectively. Evidently we also have

$$\begin{aligned} I_1(s_1) &= I_1 \\ I_1(s_2) &= 0 \end{aligned} \quad (22)$$

and

$$\int\int \bar{I}_1\cdot\bar{\bar{G}}\cdot\bar{I}_0 = 0. \quad (23)$$

In this case, the signal, V_{r1}, received at the source terminals is given by

$$V_{r1} = \frac{1}{I}\left\{\int \bar{E}\cdot\bar{I}_1\right\}. \quad (24)$$

3) The signal received from a radiating loop excited by a current I_2 at terminals 2-2 but with 1-1 open. It is clear that all the explanations and equations for 2) can be used here simply by interchanging the subscripts 1 and 2.

Now in the general case of a loaded loop, the signal V received is the sum of V_0, V_{r1}, and V_{r2}. It can be written as

$$V = V_0 + \frac{I_1}{I}\int \bar{E}\cdot\check{I}_1 + \frac{I_2}{I}\int \bar{E}\cdot\check{I}_2 \quad (25)$$

with

$$\check{I}_1(s) = \frac{\bar{I}_1(s)}{I_1} \quad \text{and} \quad \check{I}_2(s) = \frac{\bar{I}_2(s)}{I_2}. \quad (26)$$

As in the dipole case, both $\check{I}_1(s_1)$ and $\check{I}_2(s_2)$ have unit magnitude.

If the diode junction impedances at terminals 1-1 and 2-2 are denoted by Z_{L1} and Z_{L2} respectively, then the voltages V_1 and V_2 across 1-1 and 2-2 can be written as

$$\begin{aligned} V_1 &= -Z_{L1}I_1 \\ V_2 &= -Z_{L2}I_2. \end{aligned} \quad (27)$$

Using the principle of superposition, we have

$$\begin{aligned} V_1 &= V_{10} + V_{11} + V_{12} \\ V_2 &= V_{20} + V_{21} + V_{22}. \end{aligned} \quad (28)$$

Combining (18), (21), (27), and (28), the following relation is obtained:

$$\begin{aligned} &-Z_{L1}I_1\delta(s - s_1) - Z_{L2}I_2\delta(s - s_2) \\ &\quad = \bar{E}\cdot\hat{s} + \int \bar{I}_0\cdot\bar{\bar{G}}\cdot\hat{s} + \int \bar{I}_1\cdot\bar{\bar{G}}\cdot\hat{s} + \int \bar{I}_2\cdot\bar{\bar{G}}\cdot\hat{s}. \end{aligned} \quad (29)$$

Multiplying (29) by $\bar{I}_1(s)$ and integrating, we have after using $\bar{I}_1(s) = I_1\check{I}_1(s)$,

$$\begin{aligned} -Z_{L1}I_1 = &\int \bar{E}\cdot\check{I}_1 + I_1\int\int \check{I}_1\cdot\bar{\bar{G}}\cdot\check{I}_1 \\ &+ I_2\int\int \check{I}_2\cdot\bar{\bar{G}}\cdot\check{I}_1. \end{aligned} \quad (30)$$

Similarly, we also have

$$\begin{aligned} -Z_{L2}I_2 = &\int \bar{E}\cdot\check{I}_2 + I_1\int\int \check{I}_2\cdot\bar{\bar{G}}\cdot\check{I}_1 \\ &+ I_2\int\int \check{I}_2\cdot\bar{\bar{G}}\cdot\check{I}_2. \end{aligned} \quad (31)$$

The four double integrals in (30) and (31) are known as the self and mutual impedances of the loop between the two pairs of terminals. Because of the symmetry conditions assumed for the loop, we may denote the self and mutual impedances simply by Z_s and Z_m. Then we have

$$\begin{aligned} Z_s &= \int\int \check{I}_1\cdot\bar{\bar{G}}\cdot\check{I}_1 = \int\int \check{I}_2\cdot\bar{\bar{G}}\cdot\check{I}_2 \\ Z_m &= \int\int \check{I}_1\cdot\bar{\bar{G}}\cdot\check{I}_2 = \int\int \check{I}_2\cdot\bar{\bar{G}}\cdot\check{I}_1. \end{aligned} \quad (32)$$

Z_s may be described as the loop self impedance across either pair of terminals when the other pair is open, and Z_m, as the loop mutual impedance between the two pairs of terminals of the loop. In terms of Z_s and Z_m, we can then express I_1 and I_2 in (30) and (31) by the following matrix form:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = -\begin{bmatrix} Z_s+Z_{L1} & Z_m \\ Z_m & Z_s+Z_{L2} \end{bmatrix}^{-1}\begin{bmatrix} \int \overline{E}\cdot\check{I}_1 \\ \int \overline{E}\cdot\check{I}_2 \end{bmatrix}. \tag{33}$$

Using this relation, (25) can now be written as

$$V = V_0 - \frac{1}{I}\left[\int \overline{E}\cdot\check{I}_1, \int \overline{E}\cdot\check{I}_2\right] \cdot \begin{bmatrix} Z_s+Z_{L1} & Z_m \\ Z_m & Z_s+Z_{L2} \end{bmatrix}^{-1}\begin{bmatrix} \int \overline{E}\cdot\check{I}_1 \\ \int \overline{E}\cdot\check{I}_2 \end{bmatrix}. \tag{34}$$

For the purpose of applying the above relation to the measurement of $\overline{H}$ field, the following two currents—a loop current $\check{I}_l(s)$ and a dipole current $\check{I}_d(s)$—are defined as

$$\begin{aligned} \check{I}_l(s) &= \check{I}_1(s) + \check{I}_2(s) \\ \check{I}_d(s) &= \check{I}_1(s) - \check{I}_2(s). \end{aligned} \tag{35}$$

The equations in (35) imply

$$\begin{aligned} \check{I}_1(s) &= \tfrac{1}{2}(\check{I}_l + \check{I}_d) \\ \check{I}_2(s) &= \tfrac{1}{2}(\check{I}_l - \check{I}_d). \end{aligned} \tag{36}$$

Substituting (36) into (34) and simplifying, we have

$$V = V_0 - \frac{1}{4ID}\left[\int \overline{E}\cdot\check{I}_l, \int \overline{E}\cdot\check{I}_d\right] \cdot \begin{bmatrix} 2(Z_s-Z_m)+Z_{L1}+Z_{L2}, & Z_{L2}-Z_{L1} \\ Z_{L2}-Z_{L1}, & 2(Z_s+Z_m)+Z_{L1}+Z_{L2} \end{bmatrix}\begin{bmatrix} \int \overline{E}\cdot\check{I}_l \\ \int \overline{E}\cdot\check{I}_d \end{bmatrix} \tag{37}$$

where

$$D = (Z_s + Z_{L1})(Z_s + Z_{L2}) - Z_m{}^2.$$

The symmetric construction of the loop implies that the two different analytical expressions $\check{I}_1(s)$ and $\check{I}_2(s)$ really represent the same current distribution with respect to their feeding terminals. Both $\check{I}_1(s)$ and $\check{I}_2(s)$ have unit magnitude at their feeding terminals and are zero at the other pair of terminals. If the loop size is small in comparison with the wavelength, it is reasonable to assume that both $\check{I}_1(s)$ and $\check{I}_2(s)$ vary linearly from one pair of loop terminals to the other. This implies that $\check{I}_l$ is a constant circulating current of unit magnitude flowing around the loop. Using this condition and Stokes' theorem, we have

$$\int \overline{E}\cdot\check{I}_l = \int_A \nabla\times\overline{E}\cdot\hat{n}dA \tag{38}$$

where $\hat{n}$ is a unit vector normal to the loop plane and A is the area enclosed by the loop. One of the Maxwell's equations reads

$$\nabla\times\overline{E} = -j\omega\mu_0\overline{H}. \tag{39}$$

If it is combined with the assumption that $\overline{H}$ is approximately constant over the loop, we then have

$$\int \overline{E}\cdot\check{I}_l \cong -j\omega\mu_0 A H_n \tag{40}$$

where H_n is the component of $\overline{H}$ normal to the loop plane, ω is the microwave angular frequency, and μ_0 is, the permeability of free space.

At the same time, if $\overline{E}$ is also assumed to be approximately constant around the loop, then we may write

$$\int \overline{E}\cdot\check{I}_d \cong \overline{E}\cdot\int \check{I}_d. \tag{41}$$

The same symmetry conditions imply also that the vector obtained by evaluating the integral $\int\check{I}_d$ around the loop will have a direction, $\hat{d}$, which is in the loop plane and perpendicular to the line joining the two diode junctions. This means that the loop with the excitation $\check{I}_d$ acts just as a dipole along the direction $\hat{d}$. The magnitude of the integral is defined as the effective length l_d of the loop in that direction. Therefore we again obtain

$$\int \overline{E}\cdot\check{I}_d = E_d l_d \tag{42}$$

with E_d as the component of E along the direction $\hat{d}$. In fact, $\hat{d}$ is the direction of the two diodes which form two parallel but opposite sides of the loop.

Substituting (40) and (42) into (37), we have

$$\begin{aligned} V = V_0 - \frac{1}{2ID}\Bigg\{&\left(Z_s - Z_m + \frac{Z_{L1}+Z_{L2}}{2}\right)(-j\omega\mu_0 A H_n)^2 \\ &- (Z_{L1}-Z_{L2})(E_d l_d)(-j\omega\mu_0 A H_n) \\ &+ \left(Z_s + Z_m + \frac{Z_{L1}+Z_{L2}}{2}\right)(E_d l_d)^2\Bigg\}. \end{aligned} \tag{43}$$

If it is further assumed that $Z_{L1}=Z_{L2}=Z_L$, (43) is then simplified to

$$V = V_0 + K\frac{\omega^2\mu_0{}^2 A^2 H_n{}^2}{Z_s + Z_m + Z_L} - K\frac{E_d{}^2 l_d{}^2}{Z_s - Z_m + Z_L} \tag{44}$$

where $K=1/2I$ is a constant, if the source current I is kept fixed. Again Z_L is the only quantity affected by the modulation.

Some properties on the back-scattered signal V, given by (44), are observed as follows:

a) The continuous signal V_0 is not modulated, there-

fore it is not detected by the coherent detection system.

b) The second term and the third term are modulated, therefore both are subjected to detection.

c) The signal due to the second term is a maximum when $\overline{H}$ is normal to the loop plane, and it is zero when $\overline{H}$ is in the loop plane.

d) The signal due to the third term is a maximum when $\overline{E}$ is parallel to the direction of the diodes, and it is zero when $\overline{E}$ is perpendicular to the direction of the diodes.

These properties justify a procedure for measuring $\overline{H}$. It is clear that the detection system using a modulated diode loop scatterer generally detects a combined signal due to both the $\overline{H}$ field (second term) and the $\overline{E}$ field (third term). In order to eliminate $\overline{E}$ field effect, a modulated diode scatterer is used first to determine the direction of $\overline{E}$. With the direction of $\overline{E}$ known, the $\overline{H}$ field can then be determined by properly orienting the loop. As long as the direction of the diodes is kept perpendicular to $\overline{E}$, the $\overline{E}$ field effect is not detected. If the field to be measured is elliptically polarized, a modified procedure for measuring $\overline{H}$ is given in the Appendix.

The procedure given above applies to the case $Z_{L1}=Z_{L2}=Z_L$, but the same procedure also applies to the case $Z_{L1}\neq Z_{L2}$. This can be seen from (43) which gives the back-scattered signal when $Z_{L1}\neq Z_{L2}$. The relation (43) is considerably more complex than the relation (44), but it still possesses all the essential properties useful for the measurement of the $\overline{H}$ field. In other words, the equal diode junction impedances requirement may be preferred but it is not really necessary.

Appendix

A Procedure for Measuring $\overline{H}$ of an Elliptically-Polarized Field

For an elliptically-polarized field, the $\overline{E}$ vector at a fixed point always lies in a plane[5] which will be called the $\overline{E}$ plane in the following description. It is clear that all the properties of the elliptically-polarized $\overline{E}$ field can be determined by using a diode dipole scatterer, so is the $\overline{E}$ plane. With the $\overline{E}$ plane known, the following procedure or its modification may then be used to determine the $\overline{H}$ field. For the convenience of explanation, a local coordinate system associated with the point at which the $\overline{H}$ field is to be measured will be selected as follows: any two mutually-perpendicular axes in the $\overline{E}$ plane may be used as the x axis and the y axis, and the normal to the $\overline{E}$ plane as the z axis. The measurement of the $\overline{H}$ field is equivalent to the measurement of the three components H_x, H_y, H_z. These can be obtained by locating the geometrical center of the diode loop at the origin of the local coordinate system and then by using the relation (44) four times. First, by setting the diode loop such that the direction of the diodes is perpendicular to the $\overline{E}$ plane and the loop normal is parallel to the x axis, then H_x is obtained. Next, by turning the diode loop about the z axis until the loop normal is parallel to the y axis, and H_y is obtained. By further turning the loop about the y axis until the direction of the diodes is parallel to the x axis, then a combined effect of H_y and E_x is obtained. Eliminating the known result of H_y, E_x is obtained. Finally, by turning the loop about the x axis until the loop lies in the $\overline{E}$ plane, a combined effect of H_z and the same E_x is obtained. By eliminating E_x, H_z is obtained. The above procedure illustrates clearly that the modulated scattering method can be used for measuring the most general type of fields.

[5] Proof of this statement is omitted.

Acknowledgment

The author would like to thank Dr. D. K. Cheng of Syracuse University, Syracuse, N. Y., for accepting and using these scattering methods of measuring $\overline{E}$ and $\overline{H}$ distributions in a project under his direction. The measurements were first set up by J. Reale and later refined by R. Holman and B. Strait. The measured results support strongly the properties given in this paper.

Analysis of Certain Transmission-Line Networks in the Time Domain*

W. J. GETSINGER†

Summary—Many linear components in nondispersive transmission line are made up solely of commensurate lengths of line of various characteristic impedances. Such components have impulse responses that are a series of equispaced impulses, and, as a result, their frequency responses can be written as a Fourier series. Given the period and coefficients of the Fourier series describing the frequency response, the time response of the circuit to any pulse can be written down immediately as a sum of replicas of the applied pulse, each replica having an amplitude given by the coefficient of a term in the series, and occurring at a time determined by the period of that term of the series.

The pulse responses of stepped transmission-line transformers, backward-coupling hybrids, and branch-line hybrids are determined and, after assuming a simple applied-pulse shape, are plotted.

Introduction

THE use of millimicrosecond pulses requires component bandwidths that can be achieved only in the microwave frequency range. Since the usual problem is to keep pulse distortion as small as possible, nondispersive transmission lines, such as coaxial and strip transmission lines, have an advantage over dispersive lines, such as waveguide.

Many useful TEM transmission-line components contain no frequency-sensitive elements other than lengths of line. When a network has these properties, its responses to an applied pulse of any shape can be determined by making an arithmetic summation of replicas of the applied pulse, each differing from the next only in amplitude and displacement in time. This process avoids the difficulty of integrating with each applied-pulse shape separately to determine the response.

Method of Analysis

Given a network of steady-state frequency response $f(\omega)$, and an input pulse, expressed in time as $g(t)$, it is desired to find the output pulse $G(t)$ from the network.[1–4] If the network is made up entirely of commensurate lengths of nondispersive transmission lines, its frequency response $f(\omega)$ can be expanded (by methods to be described later) as an infinite series of the form

$$f(\omega) = \sum_{n=0}^{\infty} b_n e^{-in\omega T} \tag{1}$$

where b_d is a coefficient specifying magnitude, ω is radian frequency, and T is a specific time interval depending on the particular microwave component being analyzed. The Fourier integral,

$$f(t) = \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega, \tag{2}$$

where t is time, can be used to transform $f(\omega)$ into the time domain. When (2) is applied to (1), the result is

$$f(t) = \sum_{n=0}^{\infty} b_n \delta(t - nT) \tag{3}$$

where $\delta(t-nT)$ is a unit impulse occurring at time $t=nT$. This is the network impulse response, and is seen to consist of a series of impulses. The response $G(t)$ of the network to an applied pulse $g(t)$ is found by convolving $f(t)$ and $g(t)$. Convolution is described by

$$G(t) = \int_0^t g(\theta) f(t - \theta) d\theta = g(t) * f(t) \tag{4}$$

where $G(t)$ is the network response, and θ is merely a variable of integration. The lower limit is zero because the input pulse $g(t)$ is assumed to be zero before $t=0$.

Substituting (3) into (4) gives the response of the network to $g(t)$ as

$$G(t) = b_0 g(t) + b_1 g(t - T) + b_2 g(t - 2T) + \cdots \tag{5}$$

because convolution with an impulse yields a replica of the given function.

Comparison of (5) with (1) shows that if the factor T and the coefficients b_k of the frequency response of the network are known, then the response of the network to any pulse can be written down immediately as a series of replicas of the applied pulse.

The frequency response of a network is usually given in closed form, and it is necessary to expand it in order to find the pulse response. The frequency response (1) is periodic with a period $2\pi/T$, and its real part is an even function; thus, the real part of the given frequency response can be expanded in a Fourier series as

$$\mathrm{Re}\, f(\omega) = b_0 + b_1 \cos \omega T + b_2 \cos 2\omega T + \cdots \tag{6}$$

by well-known analytical or numerical methods. The coefficients b_k in (6) may be identified with those in (5), so that the pulse response of the network is known upon

* Manuscript received by the PGMTT, October 9, 1959; revised manuscript received, November 30, 1959. This work was supported by the International Business Machines Corp., Yorktown Heights, N. Y.

† Stanford Res. Inst., Menlo Park, Calif.

[1] S. Goldman, "Frequency Analysis, Modulation, and Noise," McGraw-Hill Book Co., Inc., New York, N. Y., 1st ed., pp. 125–129; 1948.

[2] M. E. Van Valkenburg, "Network Analysis," Prentice-Hall Inc., Englewood Cliffs, N. J., pp. 165, 166; 1955.

[3] E. A. Guillemin, "Communication Networks," John Wiley and Sons, Inc., New York, N. Y., vol. 2, ch. 11, pp. 461–507; 1935.

[4] E. A. Guillemin, "The Fourier integral—a basic introduction," IRE Trans. on Circuit Theory, vol. CT-2, pp. 227–230; September, 1955.

expanding only the real part of its frequency response. Real-part sufficiency is discussed extensively by Guillemin.[5]

Another method of determining the coefficients to be used in (5) is by replacing each $e^{-j\omega T}$ by ζ in the given frequency response, and expanding the function in a power series in ζ about $\zeta=0$, using any of the available methods for determining power series coefficients. The coefficients of the power series are the same as those of the Fourier series for the same values of n, as can be determined by replacing each ζ in the series with $e^{-j\omega T}$ when the expansion has been made. The substitution is possible because the networks being considered involve only commensurate lengths of nondispersive transmission line as frequency-sensitive elements, and thus the independent variable can be considered to be $(e^{-j\omega T})$ rather than ω.

Pulse Response of Stepped Transformers

A properly designed, stepped, transmission-line transformer[6] provides a means of joining two transmission lines of greatly different characteristic impedances without incurring a large mismatch. A diagram of the general step transformer to be discussed is given in Fig. 1. Assumptions made are that the transmission line is nondispersive, that junction effects can be ignored, that the physical distance between adjacent steps is the same for all steps, and that the electrical distance between adjacent steps is one-quarter wavelength at the carrier frequency of the applied pulse.

Determination of Reflected Response

The time-domain reflected-response characteristic of the step transformer will be determined by applying a unit impulse at Step 1 and obtaining the impulse response directly. Practical step transformers are usually well matched, and the reflections from the individual steps are relatively small, so that for the reflected response, signal level can be considered to be the same at each step, and multiple reflections can be ignored. Suppose a unit impulse reaches the first step at $t=0$. There is an immediate reflection at that step. If the time required for the impulse to advance from one step to the next is called T, the reflection from the kth step reaches the first step when $t=2kT$. The amplitude of each response is the voltage reflection factor ρ_k of that particular step. Thus, the reflected impulse response from all n steps can be written as

$$f_r(t) = \sum_{k=1}^{n} \rho_k \delta[t - 2(k-1)T]. \tag{7}$$

From the preceding section, the reflected response to any applied pulse can be written immediately as

$$G_r(t) = \sum_{k=1}^{n} \rho_k g[t - 2(k-1)T]. \tag{8}$$

Let

$$g(t) = e(t) \sin \omega_0 t \tag{9}$$

and

$$\omega_0 T = \pi/2 \tag{10}$$

where $e(t)$ is the time description of the pulse envelope, and ω_0 is the carrier radian frequency. Also,

$$\begin{aligned} \sin\left(\omega_0 t - \frac{\pi}{2}\right) &= -\cos \omega_0 t \\ \sin(\omega_0 t - \pi) &= -\sin \omega_0 t \\ \sin\left(\omega_0 t - \frac{3\pi}{2}\right) &= \cos \omega_0 t \\ \sin(\omega_0 t - 2\pi) &= \sin \omega_0 t. \end{aligned} \tag{11}$$

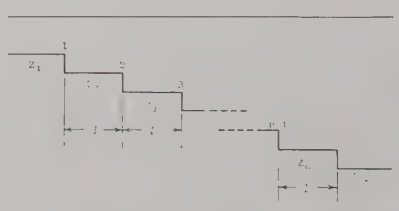


Fig. 1—The stepped transmission-line transformer.

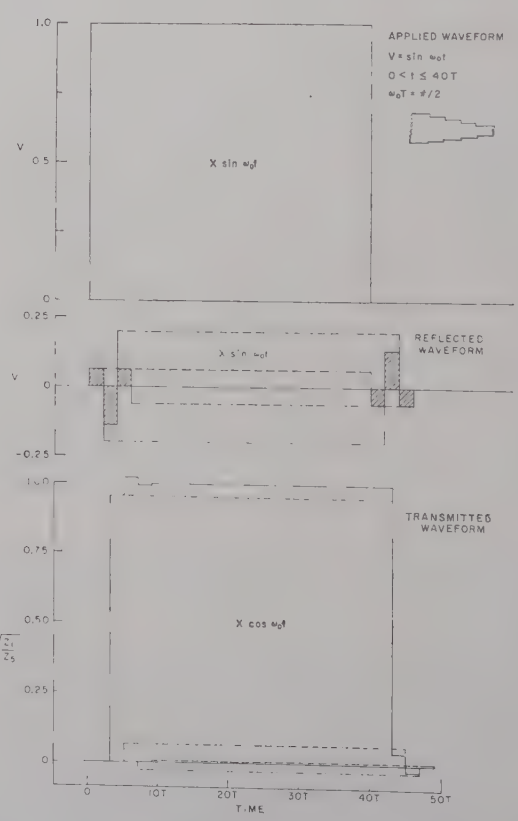


Fig. 2—Response of four-step stepped transmission-line transformer to applied-rectangular pulse.

[5] E. A. Guillemin, "Computational Techniques which Simplify the Correlation Between Steady-State and Transient Response of Filters and Other Networks," *Proc. Natl. Electronics Conf.*, Chicago, Ill., pp. 513–532; 1953.

[6] S. B. Cohn, "Optimum design of stepped transmission-line transformers," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 16–21; April, 1955.

Substitution of (9)–(11) into (8) gives

$$G_r(t) = \sin \omega_0 t \left\{ \sum_{k=1}^{n} (-1)^{k+1} \rho_k e[t - 2(k-1)T] \right\}. \quad (12)$$

Expanding for clarity, this becomes

$$G_r(t) = \sin \omega_0 t \{ \rho_1 e(t) - \rho_2 e(t - 2T) + \rho_3 e(t - 4T) - \cdots + \rho_n e[t - 2(n-1)T] \}. \quad (13)$$

The envelope of the pulse reflection response can now be plotted by addition of replicas of the applied envelope reduced in amplitude and displaced in time as given by (12) or (13). This has been done in Fig. 2 for an applied rectangular pulse $40T$ long. The dashed lines indicate the component echoes which add to give the resulting waveshape, indicated by the lined areas, for the reflected wave. Fig. 2 is for a four-step binomial transformer, but the curves do not differ appreciably from those for a four-step Tchebycheff transformer.

Determination of Transmitted Response

The time-domain transmission characteristic of the step transformer has also been determined by following a unit impulse and its significant reflections through the transformer. The unit impulse $\delta(t)$ is applied at Step 1 of Fig. 1. As the impulse travels to the right, it is changed by the voltage transmission coefficient $\vec{\tau}_k$ at each step, and sets up a reflection traveling to the left at each step. Each such reflection is partially reflected again from each step to the left of the step at which it originated. Such re-reflections travel to the right, being changed by $\vec{\tau}_k$ and setting up smaller reflections as each step is passed, and eventually emerge from the right-hand port. The multiple reflections emerging on the right are of the magnitude of $(\rho_k\rho_1)$, $(\rho_k\rho_1)^2$, $(\rho_k\rho_1)^3$, and so on. As $(\rho_k\rho_1)$ is much less than unity for practical transformers, terms of order $(\rho_k\rho_1)^2$ and greater were neglected. Also, products in $\vec{\tau}_i\overleftarrow{\tau}_i$ were approximated by unity. Assuming an applied pulse of the form $g(t) = e(t) \sin \omega_0 t$, as was done for the reflected response, the transmitted response $G_t(t)$ can be given in terms of

$$e'(t) = \left(\prod_{1}^{n} \vec{\tau}_k \right) \Bigg\{ e[t - (n-1)T] + \left(\sum_{m=2}^{n} \rho_m \rho_{m-1} \right) e[t - (n+1)T] - \left(\sum_{m=3}^{n} \rho_m \rho_{m-2} \right) e[t - (n+3)T] + \cdots + (-1)^d \left(\sum_{m=d+1}^{n} \rho_m \rho_{m-d} \right) \cdot e[t - (n + 2d - 1)T] + \text{etc.} \Bigg\} \quad (14)$$

where

$$\begin{aligned} e'(t) &= \text{the output envelope function} \\ G_t(t) &= \quad (\cos \omega_0 t) e'(t) \quad \text{for } n = 4, 8, 12, \cdots \\ G_t(t) &= -(\cos \omega_0 t) e'(t) \quad \text{for } n = 2, 6, 10, \cdots \\ G_t(t) &= \quad (\sin \omega_0 t) e'(t) \quad \text{for } n = 5, 9, 13, \cdots \\ G_t(t) &= -(\sin \omega_0 t) e'(t) \quad \text{for } n = 3, 7, 11, \cdots. \end{aligned} \quad (15)$$

Fig. 2 also shows the envelope of the transmitted response of a four-step, binomial, stepped transformer for an applied rectangular pulse. The pulse width chosen, $40T$, is equivalent to ten cycles of the carrier frequency. The response of a stepped transformer designed on other than a binomial basis would still have features similar to those of the response shown. Fig. 3 has been plotted to show the pulse response of a stepped transformer to the pulse envelope $V(t) = \sin^{1/2}(\pi t/40T)$, for $0 < t \leq 40T$. The transformer used in computing this response was a four-step Tchebycheff design with a 2:1 bandwidth ratio and 3:1 transformation ratio.

Stepped transformers, and other transmission-line components, have discontinuity reactances that are usually considered to be lumped capacitance or inductance, and thus, such a component might not be

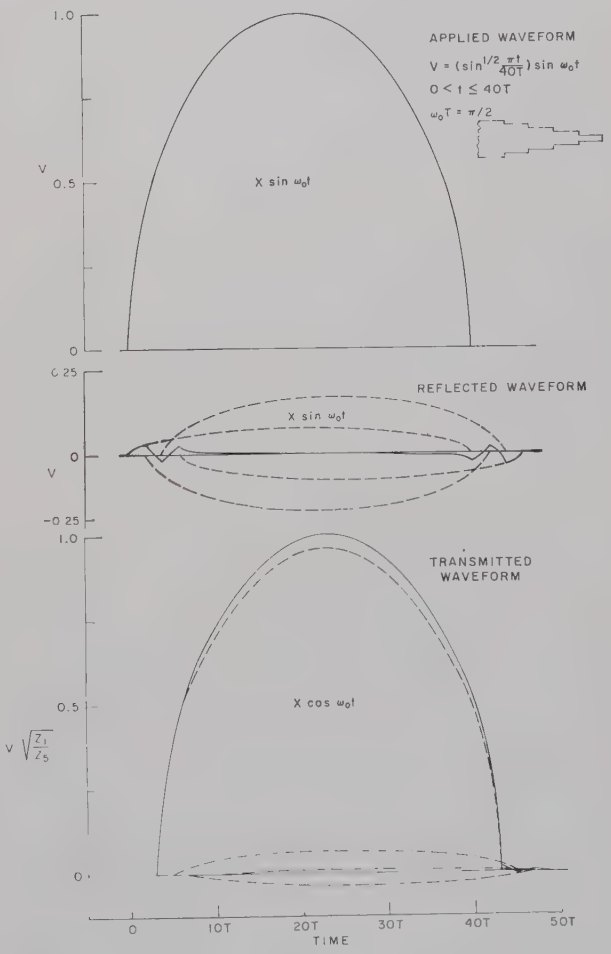


Fig. 3—Response of a four-step stepped transmission-line transformer to applied rounded pulse.

considered suited to pulse analysis by this method. However, it can be shown that reasonably small discontinuity reactances or susceptances, such as are met with in actual components, can be closely approximated by short lengths of nondispersive transmission line, if they are not so small as to be negligible. Thus, discontinuity reactances are not often a serious limitation in pulse analysis by this method.

Pulse Distortion in a Backward-Coupling Hybrid

A backward-coupling hybrid[7–10] is a symmetrical four-port network that divides RF power incident at a given port approximately equally between the port adjacent to and the port collinear with the input port, leaving the diagonally opposite port isolated, as shown in Fig. 4. The impulse responses are given in an article by Oliver,[7] and, thus, the pulse responses can be determined directly. They are:

$$\frac{V_2}{V_1}(t) = \rho g(t) - \rho(1-\rho^2)[g(t-2T) + \rho^2 g(t-4T) + \rho^4 g(t-6T) + \cdots]$$

$$\frac{V_4}{V_1}(t) = (1-\rho^2)[g(t-T) + \rho^2 g(t-3T) + \rho^4 g(t-5T) + \cdots] \quad (16)$$

for an input pulse applied at port 1 of Fig. 4. In these formulas, T is the time required for a unit impulse to traverse the length of the coupler, and ρ is defined by

$$\rho^2 = \frac{1-\sqrt{1-k^2}}{1+\sqrt{1-k^2}}, \quad (17)$$

where k is the midband amplitude coupling factor V_2/V_1. For balanced output, k has the value $\sqrt{2}/2$, and so each term in the brackets of (16) is about 15.35 db down from the preceding term. Thus, only the first few terms have significance in determining pulse distortion.

Pulse Distortion in a Backward-Coupling Hybrid Used as a Sum-and-Difference Network

To consider the backward-coupling hybrid as a sum and difference network, one input port will be Port 1′ of Fig. 4, located a distance $l=\lambda/4$ at the design-center frequency out from Port 1. This has the effect of delaying both output signals from (16) by an additional time

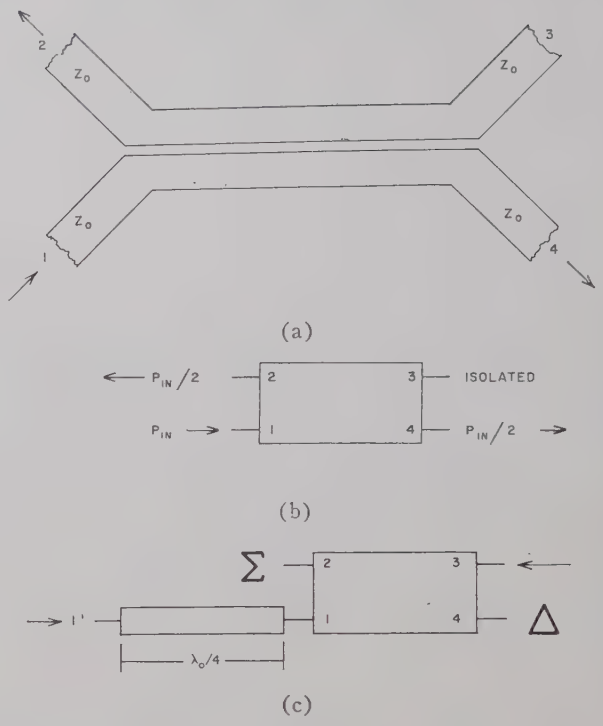


Fig. 4—The backward-coupling hybrid. (a) Physical appearance of conducting center strips. (b) Functional diagram. (c) As a sum-and-difference network.

T. Now, let identical signals, $g(t)$, be applied at Ports 1′ and 3 at the same instant. By symmetry and superposition, the total signals from Port 2, designated $H_2(t)$, and Port 4, designated $H_4(t)$, are

$$H_2(t) = \frac{V_2}{V_1}(t-T) + \frac{V_4}{V_1}(t)$$

$$H_4(t) = \frac{V_2}{V_1}(t) + \frac{V_4}{V_1}(t-T). \quad (18)$$

Solutions of (18) in terms of the parameters of (16) yield

$$H_2(t) = (1+\rho-\rho^2)g(t-T) - \rho(1-\rho)(1-\rho^2) \cdot [g(t-3T) + \rho^2 g(t-5T) + \rho^4 g(t-7T) + \cdots]$$

$$H_4(t) = \rho g(t) + (1-\rho)(1-\rho^2)[g(t-2T) + \rho^2 g(t-4T) + \rho^4 g(t-6T) + \cdots] \quad (19)$$

where $H_2(t)$ is the Σ output, and $H_4(t)$ the Δ output.

As with the stepped transformers, let $g(t)=e(t)\sin\omega_0 t$, where $e(t)$ is the function describing the envelope of the input pulse, and ω_0, the carrier frequency, is the same as the design-center frequency of the coupler. Using the trigonometric equalities of (15), (19) becomes for the sum arm

$$H_2(t) = -\cos\omega_0 t\{(1+\rho-\rho^2)e(t-T) + \rho(1-\rho)(1-\rho^2)[e(t-3T) - \rho^2 e(t-5T) + \rho^4 e(t-7t) - \cdots]\}, \quad (20)$$

[7] B. M. Oliver, "Directional electromagnetic couplers," Proc. IRE, vol. 42, pp. 1686–1692; November, 1954.

[8] S. B. Cohn, P. M. Sherk, J. K. Shimizu, and E. M. T. Jones, "Strip Transmission Lines and Components," Stanford Res. Inst., Menlo Park, Calif., Final Rept., SRI Project 1114, Contract DA 36-039 SC 63232, DA Project 3-26-00-600,SC Project 2006A; February, 1957.

[9] J. K. Shimizu, "Strip-line 3-db directional couplers," 1957 IRE WESCON Convention Record, pt. 1, pp. 4–9.

[10] J. K. Shimizu and E. M. T. Jones, "Coupled-transmission-line directional couplers," IRE Trans. on Microwave Theory and Techniques, vol. MTT-6, pp. 403–411; October, 1958.

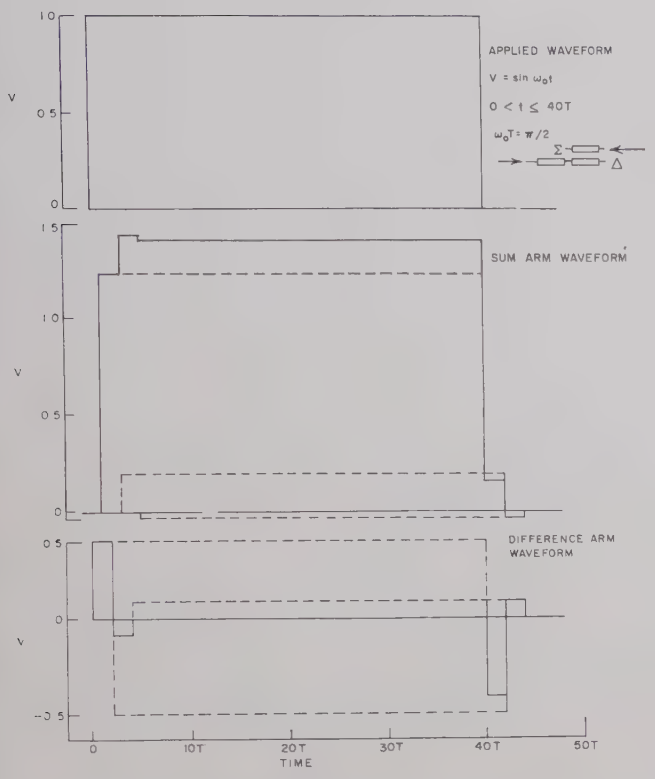


Fig. 5—Responses of backward-coupling hybrid to applied rectangular pulse.

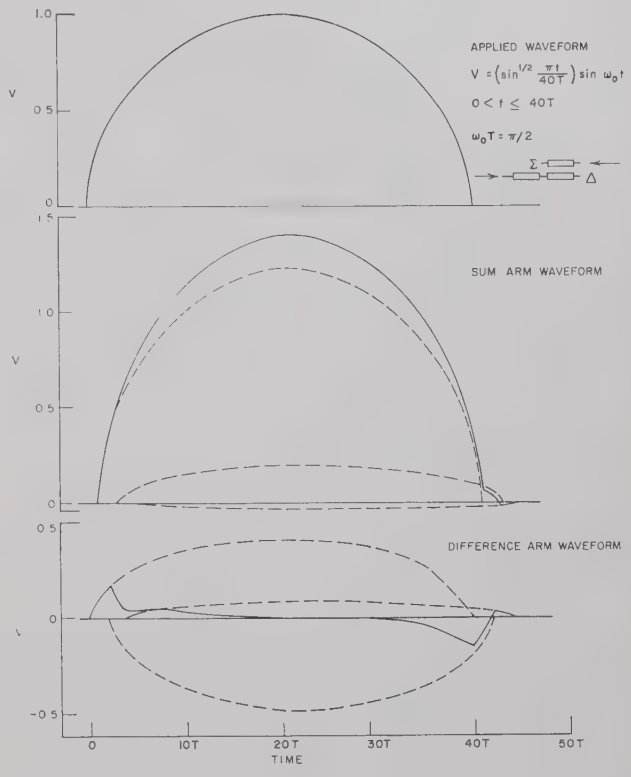


Fig. 6—Responses of backward-coupling hybrid to applied rounded pulse.

for the difference arm

$$H_4(t) = \sin \omega_0 t \{ \rho e(t) - (1-\rho)(1-\rho^2)[e(t-2T) - \rho^2 e(t-4T) + \rho^4 e(t-6T) - \cdots]\} .$$

Substitution of the numerical values $(\sqrt{2}-1)$ for ρ, appropriate for 3-db coupling, gives the pulse shapes for the sum-and-difference arms as

$$H_2(t) = -\cos \omega_0 t \{ 1.24e(t-T) + 0.20e(t-3T) - 0.03e(t-5T) \}$$

for the sum arm, and

$$H_4(t) = \sin \omega_0 t \{ 0.41e(t) - 0.49e(t-2T) + 0.08e(t-4T) \} \quad (21)$$

for the difference arm, where terms of higher order than ρ^2 are omitted, and numbers are rounded off to two places. Any desired accuracy can be obtained by using more terms of the general expression carried out to more decimal places. These equations show that, to the approximation used, all the significant echoes have made their appearance within $4T$ after the appearance of the body of the signal, and have disappeared within $4T$ after the body of the signal has ended.

To illustrate the application of (21), examples of sum-and-difference pulse responses have been plotted in Figs. 5 and 6. The $40T$ pulse width for the envelopes is equivalent to ten cycles of the carrier frequency, denoted ω_0 in (21).

The terms of (21) were added graphically. In general, for a given maximum amplitude of the pulse, the more slowly the pulse rises and falls, the smaller the maximum amplitudes of the distorting terms will be. Thus, a Gaussian-shaped pulse whose maximum slope is less than the maximum slope of the root-sine pulse described above would be expected to have a less-distorted sum-arm output and a smaller-amplitude difference-arm output.

The Branch-Line Coupler as a Sum-and-Difference Network

A coupler well suited to manufacture in strip-line is the branch-line coupler,[11] shown in Fig. 7.

In this section, the frequency response of a certain 3-db branch-line coupler will be given, along with its pulse responses. The pulse responses of the sum-and-difference ports will then be plotted for a number of 3-db branch-line couplers acting as sum-and-difference networks, allowing comparison of different designs on a pulse-response basis.

[11] J. Reed and G. J. Wheeler, "A method of analysis of symmetrical four-port networks," IRE Trans. on Microwave Theory and Techniques, vol. MTT-4, pp. 246–252; October, 1956.

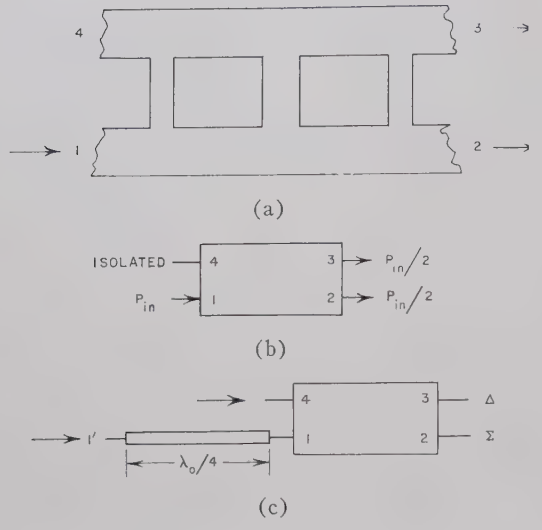


Fig. 7—The branch-line coupler. (a) Physical appearance of conducting center strip. (b) Functional diagram. (c) As a sum-and-difference network.

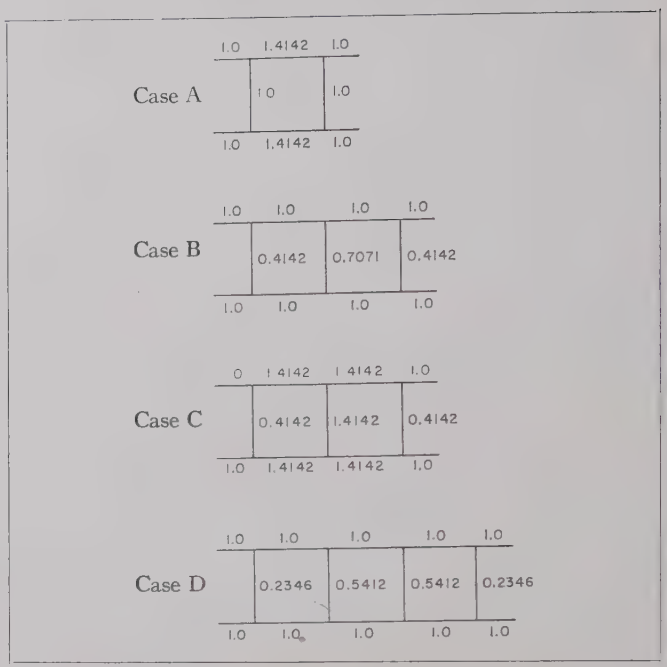


Fig. 8—Relative admittances of arms of 3-db branch-line couplers.

The admittances of the coupler arms relative to the feed-line admittance for the 3-db couplers to be considered in this report are shown in Fig. 8, where each coupler is assigned an identifying letter (A, B, C, and D, in order of increasing bandwidth). These couplers are symmetrical about their centers in all three planes. The electrical distance between adjacent cross-arms is one-quarter wavelength at center frequency, as is the length of each cross-arm. The expressions for the frequency responses of branch-line couplers are quite formidable, and are best handled by an electronic computer, such as the IBM 650.

The pulse responses of these branch-line hybrids were computed by the following method. The frequency responses between the applied signal at Port 1 and scattered signals from each of the four ports were determined by an electronic computer using the procedures of Reed and Wheeler.[11] The real parts of the frequency responses were then taken and their Fourier-series coefficients computed, also by machine. The period of the Fourier series was determined from the physical consideration that the value of the frequency response was the same at $4f_0$ as at zero frequency, where f_0 is the design-center frequency, for all ports of any of the hybrids.

The inverse of the Fourier series period is the time-delay parameter denoted T in the section on Method of Analysis. Thus,

$$T = \frac{1}{4f_0}. \tag{22}$$

This frequency period was used for all the couplers, and in each case it provided the proper time delay between signal application at Port 1 and response at any given port.

It was then necessary only to substitute in (4) to arrive at the time responses of a given coupler to an arbitrary pulse. Assuming a pulse-modulated signal of the form

$$g(t) = e(t) \sin \omega_0 t \tag{9}$$

where $e(t)$ is the envelope description and ω_0 is the design-center radian frequency of the coupler, the responses of Case B are given below for all terms of amplitude greater than 0.001 (relative to the amplitude of the applied pulse), and are plotted in Fig. 9 for a rectangular applied-pulse envelope.

$$\begin{aligned} G_1(t) = \sin \omega_0 t[&-0.172e(t) + 0.366e(t - 2T) \\ &- 0.038e(t - 4T) - 0.111e(t - 6T) \\ &- 0.046e(t - 8T) - 0.009e(t - 10T) \\ &+ 0.004e(t - 12T) + 0.004e(t - 14T) \\ &+ 0.002e(t - 16T)]. \end{aligned}$$

$$\begin{aligned} G_2(t) = -\sin \omega_0 t[&0.507e(t - 2T) + 0.068e(t - 4T) \\ &+ 0.093e(t - 6T) + 0.041e(t - 8T) \\ &+ 0.008e(t - 10T) - 0.004e(t - 12T) \\ &-0.004e(t - 14T) - 0.002e(t - 16T)]. \end{aligned}$$

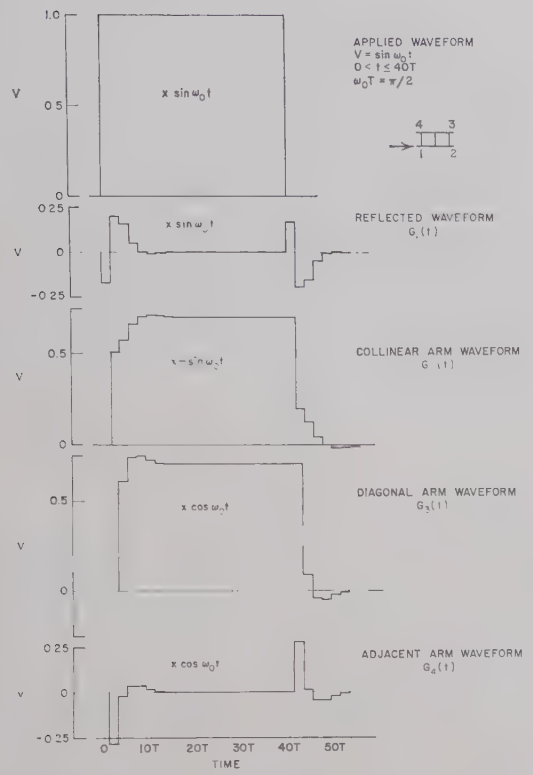


Fig. 9—Responses of 3-db branch-line coupler (Case B) to applied rectangular pulse. (Subscripts refer to port designations applied at Port 1.)

$$
\begin{aligned}
G_3(t) = \cos \omega_0 t[&0.613e(t-3T) + 0.130e(t-5T) \\
&+0.005e(t-7T) - 0.023e(t-9T) \\
&- 0.015e(t-11T) - 0.005e(t-13T)].
\end{aligned}
$$

$$
\begin{aligned}
G_4(t) = \cos \omega_0 t[&-0.284e(t-T) + 0.265e(t-3T) \\
&+ 0.060e(t-5T) - 0.001e(t-7T) \\
&- 0.022e(t-9T) - 0.014e(t-11T) \\
&- 0.005e(t-13T) - 0.000e(t-15T) \\
&+ 0.001e(t-17T)].
\end{aligned} \tag{23}
$$

These responses apply only when the applied frequency is exactly the same as the design-center frequency of the coupler. For a small change in applied frequency, the phase of the carrier of each replica is displaced by a small angle from that of the preceding replica, so that the replicas will no longer add exactly in phase. The equations of (23) can be modified to include the effect of such a change in frequency from the design-center frequency, ω_0, to a new frequency, ω_c, by removing the sin $\omega_0 t$ [or cos $\omega_0 t$] preceding the brackets and multiplying each term within the brackets by a factor sin $(\omega_c t - n\phi)$ [or cos $(\omega_c t - n\phi)$] where n is the same integer as in the term $(t - nT)$ giving the envelope displacement for that replica. The value of phase term ϕ is given by

$$\phi = \frac{\pi}{2}\left(\frac{\omega_c}{\omega_0} - 1\right). \tag{24}$$

The addition of a quarter-wavelength of line to one input port makes a sum-and-difference network of the branch-line hybrid. With reference to Fig. 7, assume that identical signals are applied at Ports 1′ and 4 simultaneously. By simple superposition, signal $H_1(t)$, emerging at Port 1′, signal $H_2(t)$, emerging at Port 2, and so on, are given by

$$
\begin{aligned}
H_1(t) &= G_1(t-2T) + G_4(t-T) \\
H_2(t) &= G_2(t-T) + G_3(t) \\
H_3(t) &= G_2(t) + G_3(t-T) \\
H_4(t) &= G_1(t) + G_4(t-T)
\end{aligned} \tag{25}
$$

because the signal takes a time interval T to travel one-quarter wavelength at the center frequency. $H_2(t)$ is the sum response and $H_3(t)$ is the difference response. The following are given for the branch-line hybrid Case B, assuming the signal to be $e(t)$ sin $\omega_0 t$, as before:

$$
\begin{aligned}
H_1(t) = \sin \omega_0 t[&-0.112e(t-2T) - 0.101e(t-4T) \\
&+ 0.098e(t-6T) + 0.110e(t-8T) \\
&+ 0.024e(t-10T) - 0.005e(t-12T) \\
&- 0.009e(t-14T) - 0.004e(t-16T) \\
&- 0.001e(t-18T)].
\end{aligned}
$$

$$
\begin{aligned}
H_2(t) = \cos \omega_0 t[&1.120e(t-3T) + 0.199e(t-5T) \\
&+ 0.097e(t-7T) + 0.018e(t-9T) \\
&- 0.006e(t-11T) - 0.008e(t-13T) \\
&- 0.004e(t-15T) - 0.002e(t-17T)].
\end{aligned}
$$

$$
\begin{aligned}
H_3(t) = \sin \omega_0 t[&-0.507e(t-2T) + 0.545e(t-4T) \\
&+ 0.038e(t-6T) - 0.036e(t-8T) \\
&- 0.032e(t-10T) - 0.011e(t-12T) \\
&- 0.001e(t-14T) + 0.002e(t-16T)].
\end{aligned}
$$

$$
\begin{aligned}
H_4(t) = \sin \omega_0 t[&-0.172e(t) + 0.082e(t-2T) \\
&+ 0.227e(t-4T) - 0.051e(t-6T) \\
&- 0.047e(t-8T) - 0.031e(t-10T) \\
&- 0.010e(t-12T) - 0.001e(t-14T) \\
&+ 0.002e(t-16T)].
\end{aligned} \tag{26}
$$

Figs. 10–13 show the sum-and-difference responses to a rectangular, pulse-modulated signal of the four branch-line couplers (Cases A, B, C, and D). It is interesting that the transients in the response curves for

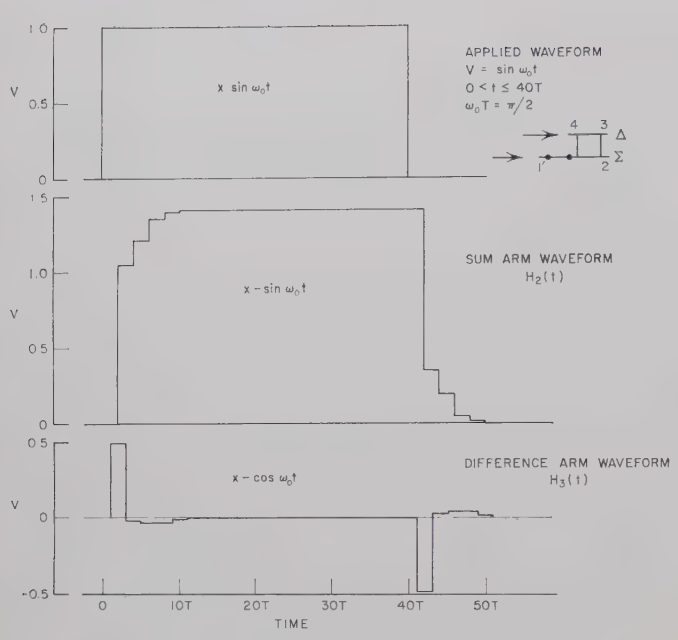


Fig. 10—Sum-and-difference responses of a branch-line coupler (Case A) to applied rectangular pulse.

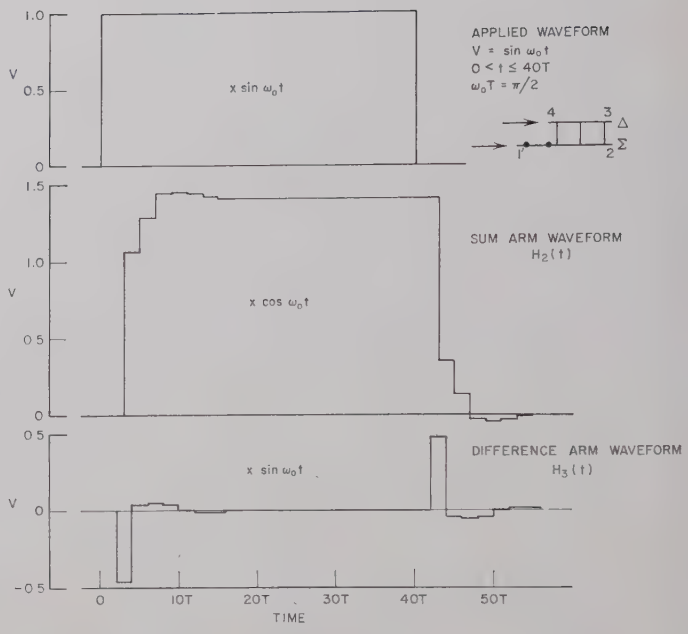


Fig. 12—Sum-and-difference responses of a branch-line coupler (Case C) to applied rectangular pulse.

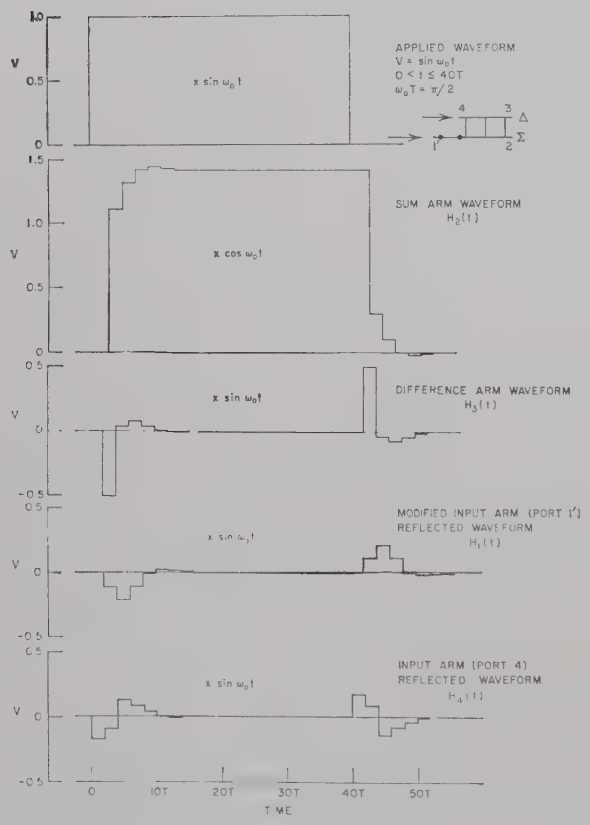


Fig. 11—Sum-and-difference responses of a branch-line coupler (Case B) to applied rectangular pulse.

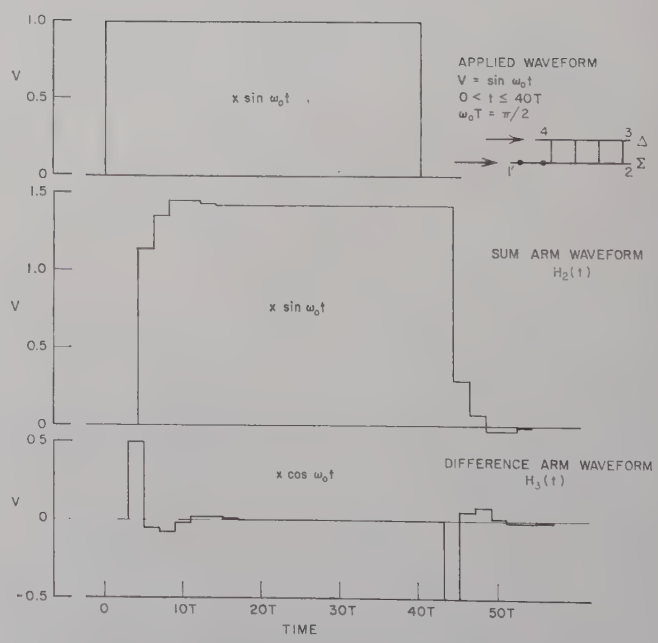


Fig. 13—Sum-and-difference responses of a branch-line coupler (Case D) to applied rectangular pulse.

the branch-line couplers and the sum-and-difference networks are of appreciable magnitude for only three cycles or less of the carrier frequency. Thus, in passing through these components, pulses only three cycles long, or spaced from each other by only three cycles, retain their general shape and identity.

Conclusion

It has been shown that the pulse responses of microwave components, made of nondispersive transmission lines only, are sums of replicas of the applied pulse. Two different ways were described by which the amplitudes and times of occurrence of the individual replicas can be found from the component frequency responses or impulse responses.

This technique for finding pulse responses was applied to stepped transmission-line transformers, to the backward coupler as a hybrid and sum-and-difference networks, and to branch-line couplers as hybrids and sum-and-difference networks. It was found that rectangular-pulse envelopes lasting for only three periods of the carrier frequency would pass through any one of these components without extreme distortion.

Acknowledgment

The author wishes to thank Dr. E. M. T. Jones for his contributions and advice, and the members of the IBM Corporation Watson Scientific Computing Laboratory for their helpful suggestions and stimulating technical discussions during the course of this work.

Sets of Eigenvectors for Volumes of Revolution*

J. VAN BLADEL†

Summary—**The electric and magnetic eigenvectors of a volume of revolution can be written in terms of two-dimensional scalar and vector functions. These functions are the eigenfunctions of certain linear transformations in the meridian plane. The form of the transformation is examined, and much attention is devoted to the orthogonality properties of their eigenfunctions and the calculation of their eigenvalues from variational principles.**

AMONG the sets of eigenvectors which exist in a finite three-dimensional volume, the "electric" and "magnetic" modes are of particular importance for the calculation of electric and magnetic fields. The purpose of the present paper is to investigate the properties of these modes in volumes of revolution of the kind depicted in Fig. 1. An explicit mathematical expression can be given for the modes of a few simple volumes, such as the sphere and the coaxial cylinder, but in the most general case one has to resort to approximate procedures to obtain quantitative data. The most frequently used methods rely on the replacement of differential equations by difference equations, and on the use of variational principles for the calculation of eigenvalues. It is necessary, for a systematic application of these methods, to possess a precise classification and enumeration of the modes and their characteristics. This is what this paper, inspired by a previous analysis by Bernier,[1] sets out to provide.

The first structure to be examined will be the toroidal volume of Fig. 1(a), which is of importance for circular particle accelerators and, more generally, for ring-like structures through which particles or fluids are flowing. The fact that a toroidal volume does not contain any portion of the axis of revolution facilitates the mathematical formulation of the problem.

* Manuscript received by the PGMTT, September 22, 1959; revised manuscript received, October 11, 1959. Research supported by the Atomic Energy Commission, Contract No. AT(11-1)-384.

† Dept. of Electrical Engineering, University of Wisconsin, Madison, Wis.

[1] J. Bernier, "On electromagnetic resonators," *Onde élect.*, vol. 26, pp. 305–317; August–September, 1946.

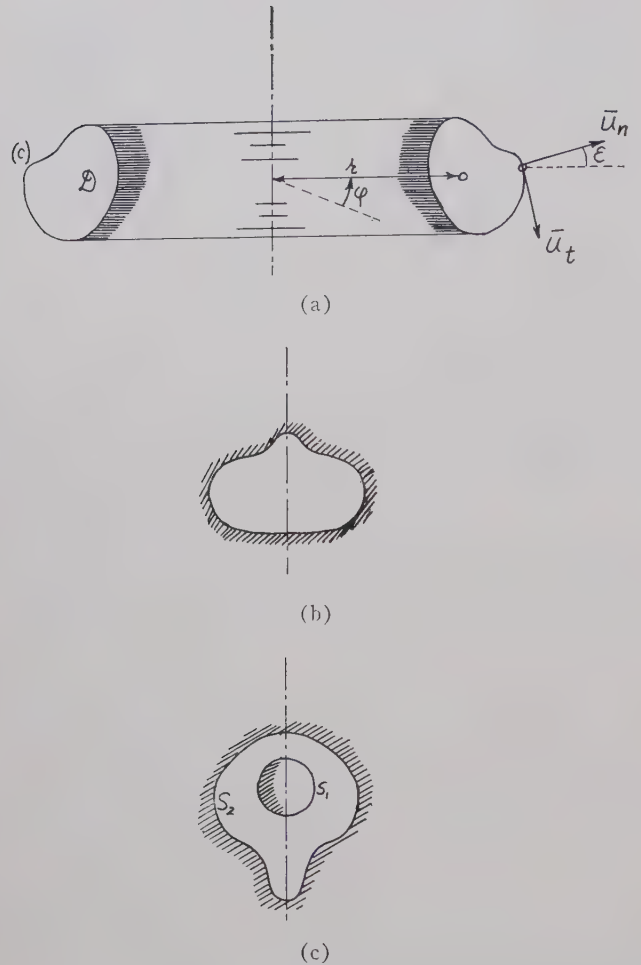


Fig. 1—Cavities of revolution.

I. Preliminary Remarks

A. Fields in Volumes of Revolution

One of the problems to be investigated is the determination of the expansion coefficients of a piecewise continuous vector function $\bar{a}(r, z, \phi)$ (r, z, and ϕ are cylindrical coordinates). This determination is simplified by a preliminary Fourier expansion in ϕ.

$$\bar{a}(r, z, \phi) = \bar{p}_0(r, z) + v_0(r, z)\bar{u}_\phi + \sum_m [\sin m\phi \cdot \bar{p}_m(r, z) + \cos m\phi \cdot \bar{q}_m(r, z)] + \sum_m [(-w_m(r, z) \sin m\phi + v_m(r, z) \cos m\phi) \cdot \bar{u}_\phi]. \quad (1)$$

The $\bar{p}$'s and $\bar{q}$'s are "meridian" vectors (*i.e.*, vectors situated in the meridian plane). Vectors such as $v_0\bar{u}_\phi$, where $\bar{u}_\phi$ is a unit vector perpendicular to the meridian plane and directed toward increasing ϕ, form the "circular" components. The divergence and curl of $\bar{a}$ are given by

$$\operatorname{div} \bar{a} = \operatorname{div}_M \bar{p}_0 + \sum_m \sin m\phi \cdot \left(\operatorname{div}_M \bar{p}_m - \frac{mv_m}{r}\right) + \sum_m \cos m\phi \cdot \left(\operatorname{div}_M \bar{q}_m - \frac{mw_m}{r}\right), \quad (2)$$

$$\operatorname{curl} \bar{a} = \operatorname{curl}_M \bar{p}_0 + \operatorname{curl} (v_0\bar{u}_\phi) + \sum_m \sin m\phi \cdot \left[\operatorname{curl}_M \bar{p}_m - \operatorname{curl}_M (w_m\bar{u}_\phi) - \frac{m}{r}(\bar{u}_\phi \times \bar{q}_m)\right] + \sum_m \cos m\phi \cdot \left[\operatorname{curl}_M \bar{q}_m + \operatorname{curl}_M (v_m\bar{u}_\phi) + \frac{m}{r}(\bar{u}_\phi \times \bar{p}_m)\right]. \quad (3)$$

Differential operators having the subscript M (M for meridian) are obtained from the usual forms by dropping derivatives with respect to ϕ and (for meridian vectors) ϕ projections.

When $\bar{a}$ is solenoidal (*i.e.*, $\operatorname{div} \bar{a} = 0$), the following relations hold:

$$\operatorname{div}_M \bar{p}_0 = 0, \quad \operatorname{div}_M \bar{p}_m = \frac{mv_m}{r}, \quad \operatorname{div}_M \bar{q}_m = \frac{mw_m}{r}. \quad (4)$$

When $\bar{a}$ is irrotational (*i.e.*, $\operatorname{curl} \bar{a} = 0$),

$$\operatorname{curl}_M \bar{p}_0 = \operatorname{curl}_M \bar{p}_m = \operatorname{curl}_M \bar{q}_m = \operatorname{curl}_M (v_0\bar{u}_\phi) = 0,$$

$$\operatorname{curl}_M (v_m\bar{u}_\phi) = -\frac{m}{r}(\bar{u}_\phi \times \bar{p}_m),$$

$$\operatorname{curl}_M (w_m\bar{u}_\phi) = -\frac{m}{r}(\bar{u}_\phi \times \bar{q}_m). \quad (5)$$

B. Electric Eigenvectors

The electric eigenvectors of a simply-bounded volume fall into two categories:

1) Irrotational eigenvectors $\bar{f}_{mnp} = \operatorname{grad} \psi_{mnp}$ where ψ_{mnp} is an eigenfunction of

$$\nabla^2\psi_{mnp} + \lambda'_{mnp}\psi_{mnp} = 0$$
$$\psi_{mnp} = 0 \text{ on boundary surface } S. \quad (6)$$

The triple index accounts for the triple infinity of eigenfunctions.

2) Solenoidal eigenvectors $\bar{e}_{mnp}$, solutions of

$$-\operatorname{curl} \operatorname{curl} \bar{e}_{mnp} + \lambda''_{mnp}\bar{e}_{mnp} = 0$$
$$\bar{u}_n \times \bar{e}_{mnp} = 0 \text{ on boundary surface } S. \quad (7)$$

The notation $\bar{u}$ stands for "unit vector," and $\bar{u}_n$ is the unit vector along the outward-pointing normal to S.

C. Magnetic Eigenvectors

The complete set of magnetic eigenvectors of a toroidal volume consists of

1) A single "sourceless" vector $\bar{h}_0 = \text{grad}\,\theta_0$, tangent to the boundary surface.[2]
2) Irrotational eigenvectors $\bar{g}_{mnp} = \text{grad}\,\theta_{mnp}$, where θ_{mnp} is an eigenfunction of

$$\nabla^2\theta_{mnp} + \nu_{mnp}\theta_{mnp} = 0 \qquad \frac{\partial\theta_{mnp}}{\partial n} = 0 \text{ on } S. \tag{8}$$

3) Solenoidal eigenvectors $\bar{h}_{mnp}$, solutions of

$$-\text{curl curl}\,\bar{h}_{mnp} + \mu_{mnp}\bar{h}_{mnp} = 0$$
$$\bar{u}_n \times \text{curl}\,\bar{h}_{mnp} = 0 \text{ on } S. \tag{9}$$

It can be shown that the eigenvalues μ and λ'' are identical, and that the electric and magnetic solenoidal eigenvectors are multiples of the curl of each other. In other words, $\bar{e}_{mnp}$ is proportional to curl $\bar{h}_{mnp}$ and $\bar{h}_{mnp}$ is proportional to curl $\bar{e}_{mnp}$. The proportionality constants depend on the normalization of the eigenvectors.

D. Variational Principle for Eigenvalues

Variational properties are of considerable interest for the approximate determination of eigenvalues and eigenvectors when the boundaries are irregular in shape. The basic property is as follows: when $\mathfrak{L}$ is a negative definite self-adjoint linear transformation,[3] all eigenvalues in $\mathfrak{L}u_n + \lambda_n u_n = 0$ are real and positive. Denoting by $\langle a, b\rangle$ the scalar product of a by b, the lowest eigenvalue λ_1 is the minimum of

$$J(u) = -\frac{\langle \mathfrak{L}u, u\rangle}{\langle u, u\rangle}.$$

This minimum is attained for the lowest eigenfunction u_1. The functions admitted for competition (the "admissible" functions) must belong to the space of definition of the transformation $\mathfrak{L}$. The second lowest eigenvalue is the minimum value of J with respect to admissible functions that are orthogonal to u_1 (*i.e.*, for which $\langle u, u_1\rangle = 0$), and the minimum is attained for $u = u_2$. Similarly, λ_n is the minimum of J with respect to u's that are orthogonal to the $(n-1)$ lowest eigenfunctions, and the minimum is attained for $u = u_n$. Similar results are obtained, *mutatis mutandis*, for positive-definite transformations.

These considerations can be applied to transformations (6) and (8). The scalar product to be used here is $\iiint_V abd\,V$, and the λ'_{mnp} are obtained as stationary values of

$$J(\psi) = -\frac{\iiint_V \psi\nabla^2\psi dV}{\iiint_V \psi^2 dV}. \tag{10}$$

The admissible functions vanish on the boundary, and are continuous up to their second derivatives. The eigenvalues ν_{mnp} are obtained as stationary values of the same expression, the admissible functions having the same continuity properties, but a vanishing normal derivative on S.

Transformation (7) with scalar product $\langle \bar{a}, \bar{b}\rangle = \iiint_V \bar{a}\cdot\bar{b}\,dV$ leads to the characterization of λ''_{mnp} as stationary value of

$$J(\bar{e}) = \frac{\iiint_V \bar{e}\cdot\text{curl curl}\,\bar{e}dV}{\iiint_V \bar{e}\cdot\bar{e}dV} \tag{11}$$

where the admissible vectors have zero divergence, are continuous up to their second derivatives, and are perpendicular to the boundary surface.

II. Electric Modes in Toroidal Volumes of Revolution

The general considerations of the preceding paragraph will now be applied more specifically to volumes of revolution.

A. Irrotational Eigenvectors

The general expression for these eigenvectors is

$$\bar{t}_{mnp} = \text{grad}[\sin m\phi\cdot\alpha_{mnp}(r, z)]$$
$$= \sin m\phi\cdot\text{grad}_M\,\alpha_{mnp} + \frac{m\cos m\phi}{r}\alpha_{mnp}\bar{u}_\phi. \tag{12}$$

The functions α are eigenfunctions of

$$\left(\nabla_M{}^2 - \frac{m^2}{r^2}\right)\alpha_{mnp} + \lambda'_{mnp}\alpha_{mnp} = 0$$

with

$$\alpha_{mnp} = 0 \text{ on } C. \tag{13}$$

Modes of revolution are obtained by setting $m = 0$ in (13). For ϕ-dependent modes, the usual ϕ degeneracy is encountered; *i.e.*, two modes, grad$[\sin m\phi\cdot\alpha]$ and grad$[\cos m\phi\cdot\alpha]$, correspond to each value of λ'. This characteristic property will be found for all other ϕ-dependent modes to be examined in the future. For reasons of conciseness, only one of the modes will be written down explicitly. The second one can then be obtained simply by increasing $m\phi$ by $\pi/2$.

[2] We define a "sourceless" vector as having zero divergence and zero curl.

[3] $\mathfrak{L}$ is self-adjoint when $\langle u, \mathfrak{L}v\rangle = \langle \mathfrak{L}u, v\rangle$ for all u, v belonging to the space of definition of $\mathfrak{L}$, and it is negative definite when $\langle u, \mathfrak{L}u\rangle \leq 0$, the equality sign being obtained for, and only for, $u = 0$. These properties are associated with a specific definition of the scalar product $\langle a, b\rangle$.

The transformation associated with (13) is self-adjoint and negative definite with respect to the scalar product $\iint_D a \cdot b \cdot r dr dz$. The α_{mnp} are orthogonal in the sense that $\iint_D \alpha_{mnp} \alpha_{mn'p'} r dr dz = 0$ for $(n, p) \neq (n', p')$. The norms of $\bar{f}$ and α are related by

$$\iiint_V |\bar{f}_{mnp}|^2 \cdot dV = \pi \iint_D \left[(\text{grad } \alpha_{mnp})^2 + \frac{m^2}{r^2} \alpha_{mnp}^2 \right] r dr dz = \lambda'_{mnp} \cdot \pi \iint_D \alpha_{mnp}^2 r dr dz. \quad (14)$$

The eigenvalues λ'_{mnp} can be obtained as stationary values of

$$J(\alpha) = - \frac{\iint_D \alpha \left[\nabla_M^2 \alpha - \frac{m^2 \alpha}{r^2} \right] r dr dz}{\iint_D \alpha^2 r dr dz}. \quad (15)$$

The admissible functions vanish on boundary (c) and have continuous derivatives up to the second order.

B. Solenoidal Eigenvectors

1) Modes of Revolution: The solenoidal eigenvectors $\bar{e}_{onp}$ can usefully be split into a meridian and a circular part according to

$$\bar{e}_{onp} = \bar{c}_{onp}(r, z) + \beta_{onp}(r, z) \bar{u}_\phi.$$

If the latter expression is substituted into (7), and the ϕ independence taken into account, uncoupled equations are obtained for $\bar{c}$ and β. The modes are consequently of two different sorts.

a) Circular modes $\beta_{onp} \bar{u}_\phi$: There is a double infinity of these modes, corresponding to the eigenfunctions of

$$\nabla_M^2 \beta_{onp} - \frac{\beta_{onp}}{r^2} + \lambda''_{onp} \beta_{onp} = 0 \quad \beta_{onp} = 0 \text{ on } C. \quad (16)$$

The β_{onp} are, in consequence, equal to the functions α_{1np} encountered in Section II, A, and partake of their orthogonality and stationarity properties. The eigenvalues λ''_{onp} of the circular modes are equal to the λ'_{1np}. The normalization is particularly simple:

$$\iiint_V \beta_{onp} \bar{u}_\phi \cdot \beta_{onp} \bar{u}_\phi dV = 2\pi \iint_D \beta_{onp}^2 r dr dz. \quad (17)$$

An example of application of variational principle (15) to calculate λ''_{onp} is given[4] in Fig. 2.

[4] For more details, see D. F. Meronek and J. Van Bladel, "Resonant modes and frequencies of a cigar-shaped cavity," *Microwave J.*, pp. 32–33; May, 1959.

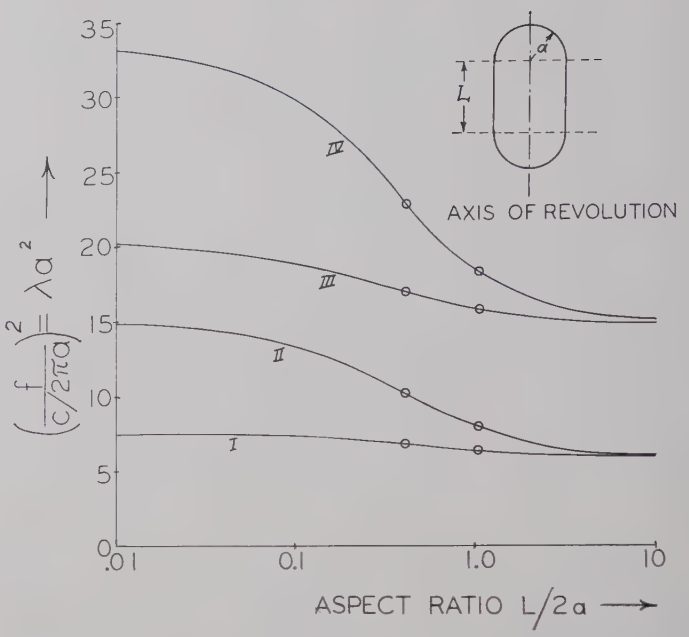


Fig. 2—Resonant frequencies of a cigar-shaped cavity. The resonant frequency is related to the eigenvalue λ^2 by $\lambda^2 = (2\pi f/c)^2$. Curves 1 and 2 correspond to the lowest two values of λ'''_{onp} (20). Curves 3 and 4 correspond to the lowest two values of λ''_{onp} (16). The circles represent experimentally determined points. (Reproduced with permission of the *Microwave Journal.*)

b) Meridian modes $\bar{c}_{onp}$: The eigenvalue problem satisfied by the $\bar{c}_{onp}$ is

$$-\text{curl curl } \bar{c}_{onp} + \lambda'''_{onp} \bar{c}_{onp} = 0 \quad \text{with} \begin{cases} \bar{u}_n \times \bar{c}_{onp} = 0 \\ \text{div}_M \bar{c}_{onp} = 0 \end{cases} \text{ on } (C). \quad (18)$$

It is a simple matter to show that these meridian vectors are actually the curl of the circular magnetic eigenvectors. More precisely, the $\bar{c}_{onp}$ can be put in the form,

$$\bar{c}_{onp} = \text{curl } [\delta_{onp} \bar{u}_\phi] = - \frac{\partial \delta_{onp}}{\partial z} \bar{u}_r + \left[\frac{\partial \delta_{onp}}{\partial r} + \frac{\delta_{onp}}{r} \right] \bar{u}_z = \frac{1}{r} [\text{grad } (\delta_{onp} r) \times \bar{u}_\phi], \quad (19)$$

where the functions δ_{onp} satisfy the eigenvalue problem,

$$\left(\nabla_M^2 - \frac{1}{r^2} \right) \delta_{onp} + \lambda'''_{onp} \delta_{onp} = 0$$

with

$$\bar{u}_n \times \text{curl } [\delta_{onp} \bar{u}_\phi] = 0 \text{ on } (C). \quad (20)$$

The boundary condition can be rewritten in the form,

$$\frac{1}{r} \frac{\partial}{\partial n} (\delta_{onp} r) = \frac{\partial \delta_{onp}}{\partial n} + \frac{\delta_{onp}}{r} \cdot \cos \epsilon = 0. \quad (21)$$

The δ_{onp} are orthogonal with respect to a scalar product $\iint_D abrdrdz$, and the eigenvalues are the stationary values of (15) with $m=1$. The admissible functions are required to possess the usual continuity properties, and to satisfy (21) at the boundary. The normalization integral is simply

$$\iiint_v \bar{c}_{onp}\cdot\bar{c}_{onp}\,dV = 2\pi\cdot\lambda'''_{onp}\iint_D \delta^2_{onp} r\,drdz. \quad (22)$$

2) Azimuth-Dependent Modes: The periodicity of the modes indicates that the general form of $\bar{e}_{mnp}$ is

$$\bar{e}_{mnp} = \bar{c}_{mnp}\sin m\phi + \bar{c}'_{mnp}\cos m\phi + [\beta_{mnp}\cos m\phi - \beta'_{mnp}\sin m\phi]\bar{u}_\phi.$$

If the latter expression is inserted in (7), uncoupled and identical equations are obtained for the pairs $(\bar{c}, \beta)$ and $(\bar{c}', \beta')$. This fact indicates the existence of an eigenvector $\bar{c}_{mnp}\sin m\phi + \beta_{mnp}\cos m\phi\bar{u}_\phi$, and also of an eigenvector $(\bar{c}', -\beta')$ obtained from the former by increasing $m\phi$ by $\pi/2$, *i.e.*, by rotating the configuration through an angle $\pi/2m$. The equations which $\bar{c}_{mnp}$ and β_{mnp} are required to satisfy are rather complicated. Dropping the subscripts for a moment, they turn out to be

$$-\operatorname{curl}_M\operatorname{curl}_M\bar{c} - \frac{m^2\bar{c}}{r^2} + \frac{m}{r}\cdot\operatorname{grad}\beta + \frac{m\beta}{r^2}\bar{u}_r + \lambda''\bar{c} = 0; \quad (23)$$

$$\nabla^2_M\beta - \frac{\beta}{r^2} + \frac{2m}{r^2}(\bar{c}\cdot\bar{u}_r) - \frac{m}{r}\operatorname{div}_M\bar{c} + \lambda''\beta = 0. \quad (24)$$

These equations can be simplified by taking into account the fact that $\bar{e}_{mnp}$ is solenoidal; *i.e.*, that

$$\operatorname{div}[\bar{c}\sin m\phi + \beta\cos m\phi\bar{u}_\phi] = \sin m\phi\left[\operatorname{div}_M\bar{c} - \frac{m\beta}{r}\right] = 0.$$

There exists, in consequence, a relation between β and $\bar{c}$, namely,

$$\beta = \frac{r}{m}\operatorname{div}_M\bar{c}. \quad (25)$$

Upon substitution of this expression in (23), an equation for $\bar{c}$ alone is obtained.

$$\nabla^2_M\bar{c}_{mnp} - \frac{m^2}{r^2}\bar{c}_{mnp} + \bar{u}_r\frac{2}{r}\operatorname{div}_M\bar{c}_{mnp} + \lambda''_{mnp}\bar{c}_{mnp} = 0$$

$$\text{with}\begin{cases}\bar{u}_n\times\bar{c}_{mnp} = 0\\ \operatorname{div}_M\bar{c}_{mnp} = 0\end{cases}\text{on }(C). \quad (26)$$

The meridian part of a solenoidal eigenvector must, in consequence, be an eigenvector of (26). Conversely, to each eigenvector of (26) corresponds an eigenvector,

$$\bar{c}_{mnp}\cdot\sin m\phi + \frac{r}{m}\operatorname{div}_M\bar{c}\cdot\cos m\phi\bar{u}_\phi, \quad (27)$$

of the original three-dimensional problem (7). It is important to list orthogonality and stationarity properties of the $\bar{c}_{mnp}$. These properties can be obtained from the general equation (9) wherein (27) is substituted. They can also be established directly from a study of the transformation,

$$\mathfrak{L}\bar{v} = \nabla^2_M\bar{v} - \frac{m^2}{r^2}\bar{v} + \frac{2\bar{u}_r}{r}\operatorname{div}_M\bar{v}$$

$$\text{with}\begin{cases}u_n\times\bar{v} = 0\\ \operatorname{div}_M\bar{v} = 0\end{cases}\text{on } C, \quad (28)$$

in the meridian plane. The relevant steps are collected in the Appendix. It turns out that the scalar product which is suited to the problem is

$$\langle\bar{v}, \bar{w}\rangle = \iint_D\left[\bar{v}\cdot\bar{w} + \frac{r^2}{m^2}\operatorname{div}_M\bar{v}\cdot\operatorname{div}_M\bar{w}\right]rdrdz \quad (29)$$

where $\bar{v}$ and $\bar{w}$ are two meridian vectors. With the latter definition of the scalar product, transformation $\mathfrak{L}$ is self-adjoint and negative-definite, the eigenvectors are orthogonal in the sense that

$$\langle\bar{c}_{mnp}, \bar{c}_{mn'p'}\rangle = \iint_D\left[\bar{c}_{mnp}\cdot\bar{c}_{mn'p'} + \frac{r^2}{m^2}\operatorname{div}_M\bar{c}_{mnp}\cdot\operatorname{div}_M\bar{c}_{mn'p'}\right]rdrdz = 0$$

$$\text{for } (n, p)\neq(n'p')$$

and the eigenvalues λ''_{mnp} are obtained from the stationarity properties of

$$J(\bar{c}) = -\frac{\langle\mathfrak{L}\bar{c}, \bar{c}\rangle}{\langle\bar{c}, \bar{c}\rangle} = -\frac{\iint_D\left[\bar{c}\cdot\mathfrak{L}\bar{c} + \frac{r^2}{m^2}\operatorname{div}_M\bar{c}\cdot\operatorname{div}_M\mathfrak{L}\bar{c}\right]rdrdz}{\iint_D\left[\bar{c}\cdot\bar{c} + \frac{r^2}{m^2}\operatorname{div}_M\bar{c}\cdot\operatorname{div}_M\bar{c}\right]rdrdz} \quad (30)$$

where the $\bar{c}$ have continuous derivatives up to the second, and satisfy the boundary conditions evidenced in (28). Third order derivatives appear in the numerator.

An equivalent expression for the latter can be derived which involves lesser order derivatives only. The derivation is based on

$$\iint_D \left\{ \bar{c}\cdot\mathfrak{L}\bar{c} + \frac{r^2}{m^2}\operatorname{div}_M \bar{c}\operatorname{div}_M \mathfrak{L}\bar{c} + \left|\operatorname{curl}_M \bar{c}\right|^2 + \left[\frac{m}{r}\bar{u}_\phi \times \bar{c} + \operatorname{curl}_M\left(\frac{r}{m}\operatorname{div}_M \bar{c}\right)\bar{u}_\phi\right]^2\right\} r\,dr\,dz$$

$$= \int_c \left\{(\bar{u}_n \times \bar{c})\cdot\operatorname{curl}_M \bar{c} - \operatorname{div}_M \bar{c}(\bar{c}\cdot\bar{u}_n) - \frac{r}{m}\operatorname{div}_M \bar{c}\left[\bar{u}_t\cdot\operatorname{curl}_M\left(\frac{r}{m}\operatorname{div}_M \bar{c}\right)\bar{u}_\phi\right]\right\} r\,dc, \quad (31)$$

a direct consequence of the substitution of (27) in the general relation,

$$\iiint_v \left[-\bar{v}\cdot\operatorname{curl}\operatorname{curl}\bar{v} + \operatorname{curl}\bar{v}\cdot\operatorname{curl}\bar{v}\right] dv = \iint_S (\bar{v}\times\operatorname{curl}\bar{v})\cdot\bar{u}_n dS. \quad (32)$$

The right-hand member of (31) vanishes for all admissible vectors. As a consequence, $J(\bar{c})$ can be rewritten as $J(\bar{c}) = N/D$ with

$$N = \iint_D \left\{\left[\frac{mc_z}{r} - \frac{r}{m}\frac{\partial^2 c_r}{\partial r\partial z} - \frac{1}{m}\frac{\partial c_r}{\partial z} - \frac{r}{m}\frac{\partial^2 c_z}{\partial z^2}\right]^2 + \left[\frac{\partial c_r}{\partial z} - \frac{\partial c_z}{\partial r}\right]^2 + \left[\frac{r}{m}\frac{\partial^2 c_r}{\partial r^2} + \frac{3}{m}\cdot\frac{\partial c_r}{\partial r} + \frac{2}{m}\cdot\frac{\partial c_z}{\partial z} + \frac{r}{m}\cdot\frac{\partial^2 c_z}{\partial r\partial z} + \frac{c_r}{r}\left(\frac{1}{m} - m\right)\right]^2\right\} r\,dr\,dz$$

$$D = \iint_D \left[c_r^2 + c_z^2 + \frac{r^2}{m^2}\left(\frac{\partial c_r}{\partial r} + \frac{c_r}{r} + \frac{\partial c_z}{\partial z}\right)^2\right] r\,dr\,dz.$$

This form is suitable for numerical computations. We repeat that the admissible vectors must satisfy the conditions,

$$\operatorname{div}_M \bar{c} = \frac{\partial c_r}{\partial r} + \frac{c_r}{r} + \frac{\partial c_z}{\partial z} = 0,$$

$$\left|\bar{c}\times\bar{u}_n\right| = c_r\sin\epsilon - c_z\cos\epsilon = 0, \quad (33)$$

at the boundary.

Finally, the normalization relations are

$$\iiint_V \bar{c}_{mnp}\cdot\bar{c}_{mnp}\,dV = \pi\iint_D \left[\bar{c}_{mnp}\cdot\bar{c}_{mnp} + \frac{r^2}{m^2}(\operatorname{div}_M \bar{c}_{mnp})^2\right] r\,dr\,dz. \quad (34)$$

III. Expansion in Electric Eigenvectors

We now turn to the task of determining the coefficients of expansion of the vector function $\bar{a}$ considered in the first paragraph.[5] The expansion breaks down in separate expansions involving the various Fourier coefficients:

$$\begin{aligned}
\bar{p}_0(r,z) &= \sum_n\sum_p A_{onp}\operatorname{grad}\alpha_{onp} + \sum_n\sum_p D_{onp}\bar{c}_{onp},\\
v_0(r,z) &= \sum_n\sum_p C_{onp}\beta_{onp},\\
\bar{p}_m(r,z) &= \sum_n\sum_p A_{mnp}\operatorname{grad}\alpha_{mnp} + \sum_n\sum_p E_{mnp}\bar{c}_{mnp},\\
\bar{q}_m(r,z) &= \sum_n\sum_p B_{mnp}\operatorname{grad}\alpha_{mnp} + \sum_n\sum_p F_{mnp}\bar{c}_{mnp},\\
v_m(r,z) &= \sum_n\sum_p A_{mnp}\frac{m}{r}\alpha_{mnp} + \sum_n\sum_p E_{mnp}\frac{r}{m}\operatorname{div}_M \bar{c}_{mnp},\\
w_m(r,z) &= \sum_n\sum_p B_{mnp}\frac{m}{r}\alpha_{mnp} + \sum_n\sum_p F_{mnp}\frac{r}{m}\operatorname{div}_M \bar{c}_{mnp}.
\end{aligned} \quad (35)$$

The value of the coefficients can be calculated from (1) and (35). Results only will be quoted. For the irrotational terms:

$$A_{onp} = \frac{\iint_D \bar{p}_o\cdot\operatorname{grad}\alpha_{onp}\,r\,dr\,dz}{\lambda'_{onp}\iint_D \alpha^2_{onp}\,r\,dr\,dz} = -\frac{\iint_D \alpha_{onp}\operatorname{div}_M \bar{p}_o\,r\,dr\,dz}{\lambda'_{onp}\iint_D \alpha^2_{onp}\,r\,dr\,dz}, \quad (36)$$

$$A_{mnp} = \frac{\iint_D \left[\bar{p}_m\cdot\operatorname{grad}\alpha_{mnp} + v_m\cdot\frac{m}{r}\cdot\alpha_{mnp}\right] r\,dr\,dz}{\lambda'_{mnp}\iint_D \alpha^2_{mnp}\,r\,dr\,dz} = -\frac{\iint_D \left[\operatorname{div}_M \bar{p}_m - \frac{m}{r}v_m\right]\alpha_{mnp}\,r\,dr\,dz}{\lambda'_{mnp}\iint_D \alpha^2_{mnp}\,r\,dr\,dz}. \quad (37)$$

A similar expression can be obtained for B_{mnp} by substituting $\bar{q}_m$ and w_m for $\bar{p}_m$ and v_m, respectively. Formulas (2), (36), and (37) indicate that coefficients A and B vanish when $\bar{a}$ is solenoidal.

[5] Simpler formulas are obtained when $\bar{a}$ is specialized to be an electric or a magnetic field. This specialization will be considered in a subsequent paper where the application to particle accelerator problems will be emphasized.

For the solenoidal terms:

$$C_{onp} = \frac{\iint_D v_o \cdot \beta_{onp} r dr dz}{\iint_D \beta_{onp}^2 r dr dz} = \frac{\iint_D \operatorname{curl}_M (\beta_{onp}\bar{u}_\phi)\cdot \operatorname{curl}(v_o \bar{u}_\phi) r dr dz - \int_c v_o[\bar{u}_t \cdot \operatorname{curl}(\beta_{onp}\bar{u}_\phi)] r dc}{\lambda'''_{onp} \iint_D \beta_{onp}^2 r dr dz},$$

$$D_{onp} = \frac{\iint_D \bar{p}_o \cdot \bar{c}_{onp} r dr dz}{\iint_D |\bar{c}_{onp}|^2 r dr dz} = \frac{\iint_D \operatorname{curl}_M \bar{p}_o \cdot \operatorname{curl}_M \bar{c}_{onp} r dr dz - \int_c (\bar{u}_n \times \bar{p}_o)\cdot \operatorname{curl}_M \bar{c}_{onp} r dc}{\lambda''_{onp} \iint_D |\bar{c}_{onp}|^2 r dr dz}$$

$$= \frac{\iint_D \delta_{onp}(\operatorname{curl} \bar{p}_o \cdot \bar{u}_\phi) r dr dz - \int_c \delta_{onp}(\bar{p}_0 \cdot \bar{u}_t) r dc}{\iint_D \delta_{onp}^2 r dr dz},$$

$$E_{mnp} = \frac{\iint_D \left[\bar{p}_m \cdot \bar{c}_{mnp} + \frac{r}{m} v_m \operatorname{div}_M \bar{c}_{mnp}\right] r dr dz \cdot}{\iint_D \left[|\bar{c}_{mnp}|^2 + \frac{r^2}{m^2}(\operatorname{div}_M \bar{c}_{mnp})^2\right] r dr dz}. \tag{38}$$

The numerator can be rewritten as

$$\frac{1}{\lambda''_{mnp}} \iint_D \Big\{ \operatorname{curl}_M \bar{p}_m \cdot \operatorname{curl}_M \bar{c}_{mnp} + \left[\operatorname{curl}(v_m \bar{u}_\phi) + \frac{m}{r}(\bar{u}_\phi \times \bar{p}_m)\right] \cdot \left[\operatorname{curl}\left(\bar{u}_\phi \frac{r}{m} \operatorname{div}_M \bar{c}_{mnp}\right) + \frac{m}{r}(\bar{u}_\phi \times \bar{c}_{mnp})\right] r dr dz.$$

A similar expression can be obtained for F_{mnp} by substituting $\bar{q}_m$ and w_m for $\bar{p}_m$ and v_m, respectively. It will be noticed that the $C, D, E,$ and F vanish when $\bar{a}$ is irrotational [which, according to (3), entails vanishing of the surface integral in the numerator], and perpendicular to the boundary (which entails $\bar{u}_n \times \bar{p}_m = 0$ and $v_m = 0$, *i.e.*, vanishing of the line integral in the numerator).

IV. Magnetic Modes in Toroidal Volumes of Revolution

The complete set of eigenvectors includes, first of all, a "sourceless" vector $\bar{u}_\phi/r$. It includes, in addition, a triple infinity of irrotational eigenvectors and a triple infinity of solenoidal eigenvectors. These we now proceed to investigate.

A. Irrotational Eigenvectors

The irrotational eigenvectors are of the form $\bar{g}_{mnp} = \operatorname{grad}[\sin m\phi \cdot \gamma_{mnp}(r, z)]$ where the γ_{mnp} are eigenfunctions of (13), but with the boundary condition $\partial\gamma_{mnp}/\partial n = 0$ on c. All properties of the α_{mnp} (orthogonality, norm, etc.) are still valid provided γ and ν are substituted for α and λ', respectively. The eigenvalues can be obtained from (15), but the admissible functions are now required to have zero normal derivative on (c).

B. Solenoidal Eigenvectors

1) Modes of Revolution: Two categories of modes will be recognized here.

a) Circular modes $\delta_{onp}\bar{u}_\phi$: There is a double infinity of these modes, corresponding to the eigenfunctions of (20) with accompanying boundary conditions. The normalization relation is (17), with β replaced by δ.

b) Meridian modes $\bar{d}_{onp}$: These eigenvectors are actually the curl of the circular electric eigenvectors. In mathematical form,

$$\bar{d}_{onp} = \operatorname{curl}[\beta_{onp}\bar{u}_\phi] = \frac{1}{r}[\operatorname{grad}(r\beta_{onp}) \times \bar{u}_\phi]. \tag{39}$$

It is a simple matter to check that $\operatorname{curl} \bar{d}_{onp} = \lambda''_{onp}\beta_{onp}\bar{u}_\phi$. In consequence, $\operatorname{curl} \bar{d}_{onp}$ vanishes on the surface of the torus, and the boundary condition $\bar{u}_n \times \operatorname{curl} \bar{d}_{onp} = 0$ is satisfied there, as it should be. The normalization integral connecting $\bar{d}$ to β is similar to (22) with $\bar{c}$ and δ replaced by $\bar{d}$ and β.

2) Azimuth-Dependent Modes: The magnetic vectors are the curl of the electric vectors $\bar{e}_{mnp}$. More precisely, with $\bar{e}_{mnp}$ given by (27), $\bar{h}_{mnp}$ will be

$$h_{mnp} = \underbrace{\cos m\phi \left[\operatorname{curl}\left(\bar{u}_\phi \frac{r}{m} \operatorname{div}_M \bar{c}_{mnp}\right) + \frac{m}{r}(\bar{u}_\phi \times c_{mnp})\right]}_{\text{meridian part}} + \underbrace{\sin m\phi \operatorname{curl} \bar{c}_{mnp}}_{\text{circular part}}. \tag{40}$$

The norm of $\bar{h}$ can be evaluated with the help of (31), as:

$$\iiint_V \bar{h}_{mnp}\cdot\bar{h}_{mnp}dV = \lambda''_{mnp}\iiint \bar{e}_{mnp}\cdot\bar{e}_{mnp}dV$$

$$= \lambda''_{mnp}\pi\iint_D\left[\bar{c}_{mnp}\cdot\bar{c}_{mnp}+\frac{r^2}{m^2}(\mathrm{div}_M\,\bar{c}_{mnp})^2\right]rdrdz.$$

If $\bar{e}_{mnp}$ has been previously normalized, the normalized magnetic eigenvector is curl $\bar{e}_{mnp}/(\lambda''_{mnp})^{1/2}$. It is sometimes desirable to calculate $\bar{h}_{mnp}$ directly without relying on a previous knowledge of $\bar{e}_{mnp}$. The relevant steps are as follows:

The value of the coefficients can be calculated from (1) and (35). Results are, for the sourceless vector,

$$A_0 = \frac{\iint_D v_0 drdz}{\iint_D \dfrac{drdz}{r}}. \tag{44}$$

$v_0(r, z)$ is $1/(2\pi r)$ times the circulation of $\bar{a}$ around the "parallel" circle through r, z. For a vector which is irrotational in the toroidal region, the circulation is constant and equal to $2\pi A_0$.

For the irrotational vectors, the coefficients are

$$A_{onp} = \frac{\int_c \gamma_{onp}(\bar{u}_n\cdot\bar{p}_0)rdc - \iint_D \gamma_{onp}\,\mathrm{div}_M\,\bar{p}_0 rdrdz}{\nu_{onp}\iint_D \gamma^2_{onp}rdrdz},$$

$$A_{mnp} = \frac{\int_c \gamma_{mnp}(\bar{u}_n\cdot\bar{p}_m)rdc - \iint_D \gamma_{mnp}\left(\mathrm{div}_M\,\bar{p}_m - \dfrac{m}{r}v_m\right)rdrdz}{\nu_{mnp}\iint_D \gamma^2_{mnp}rdrdz}. \tag{45}$$

1) $\bar{h}_{mnp}$ will be of the form

$$\bar{h}_{mnp} = \bar{d}_{mnp}\sin m\phi + \frac{r}{m}\cos m\phi\bar{u}_\phi\,\mathrm{div}_M\,\bar{d}_{mnp}. \tag{41}$$

2) If we go through the same motions as with the electric eigenvectors, we discover that the meridian part $\bar{d}_{mnp}$ is an eigenvector of

$$\mathfrak{L}'\bar{d} = \nabla^2_M\bar{d}_{mnp} - \frac{m^2}{r^2}\bar{d}_{mnp} + \frac{2\bar{u}_r}{r}\mathrm{div}_M\,\bar{d}_{mnp} + \lambda'_{mnp}\bar{d}_{mnp} = 0$$

$$\text{with}\begin{cases}\bar{u}_n\cdot\bar{d} = 0\\ \mathrm{curl}_M\,\bar{d} = 0\end{cases}\text{on }(C). \tag{42}$$

Transformation $\mathfrak{L}'$ is again self-adjoint and negative-definite with respect to scalar product (29). Its eigenvectors are orthogonal, and its eigenvalues are the stationary values of (30). The admissible vectors, however, must now satisfy the boundary conditions evidenced in (42). These can be written more explicitly as

$$\begin{cases}\bar{u}_n\cdot\bar{d} = d_r\cos\epsilon + d_z\sin\epsilon = 0\\ |\,\mathrm{curl}_M\,\bar{d}\,| = \dfrac{\partial d_r}{\partial z} - \dfrac{\partial d_z}{\partial r} = 0.\end{cases}\text{on }(C), \tag{43}$$

V. Expansion in Magnetic Eigenvectors

This expansion is particularly suitable when vector $\bar{a}$ is tangent to the boundary surface. The expansion is similar to (35), with γ and $\bar{d}$ replacing α and $\bar{c}$. The only difference occurs in the expansion for v_0, which is now

$$v_0(r, z) = \frac{A_0}{r} + \sum_n\sum_p c_{onp}\,\delta_{onp}(r, z).$$

A similar expression can be obtained for B_{mnp} by substituting $\bar{q}_m$ and w_m for $\bar{p}_m$ and v_m, respectively. Formulas (2) and (45) indicate that A_{onp}, A_{mnp}, and B_{mnp} vanish when $\bar{a}$ is solenoidal and perpendicular to the boundary.

For the solenoidal vectors,

$$C_{onp} = \frac{\iint_D v_0\delta_{onp}rdrdz}{\iint_D \delta^2_{onp}rdrdz}$$

$$= \frac{\iint_D \mathrm{curl}\,(\delta_{onp}\bar{u}_\phi)\cdot\mathrm{curl}\,(v_0\bar{u}_\phi)rdrdz}{\lambda'''_{onp}\iint_D \delta^2_{onp}rdrdz},$$

$$D_{onp} = \frac{\iint_D \bar{p}_0\cdot\bar{d}_{onp}rdrdz}{\iint_D \bar{d}_{onp}\cdot\bar{d}_{onp}rdrdz}$$

$$= \frac{\iint_D \mathrm{curl}_M\,\bar{p}_0\cdot\mathrm{curl}_M\,\bar{d}_{onp}rdrdz}{\lambda''_{onp}\iint_D \bar{d}_{onp}\cdot\bar{d}_{onp}rdrdz}$$

$$= \frac{\iint_D \beta_{onp}(\mathrm{curl}\,\bar{p}_o\cdot\bar{u}_\phi)rdrdz}{\iint_D \beta^2_{onp}rdrdz},$$

$$E_{mnp} = \frac{\int\int_D \left[\bar{p}_m \cdot \bar{d}_{mnp} + \frac{r}{m} v_m \operatorname{div}_M \bar{d}_{mnp} \right] r dr dz}{\int\int_D \left[\bar{d}_{mnp} \cdot \bar{d}_{mnp} + \frac{r^2}{m^2} (\operatorname{div}_M \bar{d}_{mnp})^2 \right] r dr dz}. \quad (46)$$

The numerator can be rewritten as

$$\frac{1}{\lambda''_{mnp}} \int\int_D \Big\{ \operatorname{curl}_M \bar{p}_m \cdot \operatorname{curl}_M \bar{d}_{mnp} + \left[\operatorname{curl}_M (v_m \bar{u}_\phi) + \frac{m}{r} (\bar{u}_\phi \times \bar{p}_m) \right] \cdot \left[\operatorname{curl}_M \left(\frac{r}{m} \bar{u}_\phi \operatorname{div}_M \bar{d}_{mnp} \right) + \frac{m}{r} (\bar{u}_\phi \times \bar{d}_{mnp}) \right] \Big\} r dr dz.$$

A similar expression can be obtained for F_{mnp} by substituting $\bar{q}_m$ and w_m for $\bar{p}_m$ and v_m, respectively. It will be noticed that, according to (3), the C, D, E, and F vanish when $\bar{a}$ is irrotational.

VI. Regions Containing the Axis

In regions of the type depicted in Fig. 1(b) and 1(c), which contain parts of the axis of revolution, the Fourier expansion coefficients of a continuous function $A(r, \bar{v}, \phi)$ behave in an interesting way in the vicinity of the axis. Let the expansion be written as

$$A(r, z, \phi) = A_0(r, z) + \sum_{m=1} \sin m\phi A_m(r, z) + \sum_{m=1} \cos m\phi B_m(r, z). \quad (47)$$

If A is continuous at all points, including those situated on the axis, the limit of A as r approaches zero must be independent of ϕ. This clearly requires A_m and B_m to vanish on the axis, while the value of A reduces to $A_0(o, z)$ thereon.

Consider now a vector $\bar{a}$, continuous at all points, including those situated on the axis, and possessing a Fourier expansion of the type given in (1). By a series of simple calculations, the details of which are left to the reader, it is possible to establish the following properties of the Fourier coefficients:

1) $\bar{p}_0$ is directed along the axis;
2) v_0 vanishes on the axis;
3) $\bar{p}_1$ and $\bar{q}_1$ are purely radial on the axis, and the equalities $p_{1r} = v_1$, $q_{1r} = w_1$ hold there;
4) The coefficients $\bar{p}_m$, $\bar{q}_m$, v_m, and w_m vanish on the axis when m is larger than one.

These simple rules for scalar and vector functions allow one to foresee the behavior of functions possessing higher orders of continuity. The scalar and vector eigenelements of a cavity have continuous Laplacian and "curl curl" on the axis. Their behavior is governed by the following rules which are of great importance for practical computations:

In a simply-bounded, simply-connected cavity of the type shown in Fig. 1(b),

1) Electric and magnetic irrotational eigenvectors:
   a) When of revolution, satisfy unchanged boundary conditions on the outer contour, but the additional condition $\partial\alpha/\partial r = \partial\gamma/\partial r = 0$ on the axis.
   b) When azimuth-dependent [as described in (13)], satisfy the additional condition $\alpha = \gamma = 0$ on the axis.

2) Electric and magnetic solenoidal eigenvectors:
   a) When of the "circular mode of revolution" type [as described in (16) and (20)] satisfy the additional condition $\beta = \gamma = 0$ on the axis.
   b) When of the "meridian mode of revolution" type [as described in (18) and (39)] satisfy the additional condition $c_r = \partial c_z/\partial r = 0$ on the axis.
   c) When azimuth-dependent [as described in (26) and (42)], satisfy the additional conditions,

$$c_z = \frac{\partial c_r}{\partial r} = 0 \qquad \text{for } m = 1,$$

$$c_z = c_r = 0 \qquad \text{for } m > 1 \text{ on the axis.} \quad (48)$$

These various relations can be checked on the normal modes of the circular cylinder, which can be written down explicitly by separation of variables.[6] (See Fig 3.) The irrotational electric eigenvectors, for example, derive from scalar functions

$$\alpha_{mnp} = \sin \frac{n\pi z}{L} J_m\left(\mu_{mp} \frac{r}{a}\right)$$

where the μ_{mp} are roots of $J_m(x) = 0$. The power expansion of Bessel's function,

$$J_m(\lambda r) = \frac{(\lambda r)^m}{2^m m!} \left[1 - \frac{(\lambda r)^2}{4(m+1)} + \cdots \right] (m \text{ integer}),$$

confirms that $\partial\alpha_{onp}/\partial r = \alpha_{mnp} = 0$ on the axis. Another check is afforded by the expression for the solenoidal ϕ-dependent electric eigenvectors:

$$\bar{c}_{mnp} = -\frac{n\pi}{L} \cdot \frac{a^2}{\mu_{mp}^2} \cdot \sin \frac{n\pi z}{L} \cdot \frac{dJ_m\left(\mu_{mp} \frac{r}{a}\right)}{dr} \bar{u}_r + \cos \frac{n\pi z}{L} J_m\left(\mu_{mp} \frac{r}{a}\right) \bar{u}_z.$$

[6] See, *e.g.*, C. G. Montgomery "Techniques of Microwave Measurements," McGraw-Hill Book Co., Inc., New York, N. Y., p. 297; 1947.

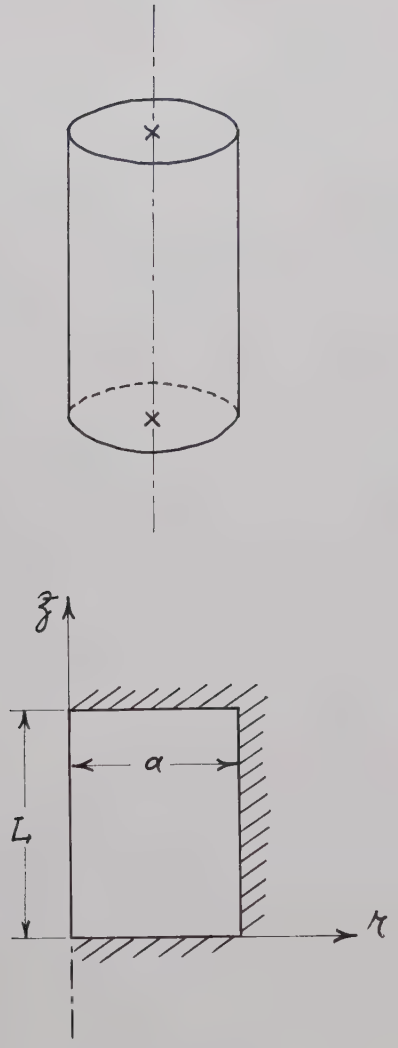


Fig. 3—Circular cylindrical cavity.

The (r, z) dependent part of the circular component is

$$v_m(r, z) = \frac{r}{m} \operatorname{div}_M \bar{c}_{mnp}$$

$$= -\frac{n\pi}{L} \cdot \frac{m}{r} \cdot \frac{a^2}{\mu_{mp}^2} \cdot \sin \frac{n\pi z}{L} \cdot J_m\left(\mu_{mp} \frac{r}{a}\right).$$

It is immediately apparent that the relevant conditions (48) are satisfied.

A last remark is in order concerning doubly-bounded volumes of the kind represented in Fig. 1(c). The electric eigenvectors considered up to now do not form a complete set unless we add the electrostatic field grad α_0 to them. This field is obtained by establishing a potential difference between the two boundary surfaces, assumed to be metallized. More precisely, α_0 is the solution of

$$\nabla_M^2 \alpha_0 = 0,$$

$$\alpha_0 = 1 \text{ on } S_1, \qquad \alpha_0 = 0 \text{ on } S_2 \tag{49}$$

(or any multiple thereof).

Appendix

Properties of the Operator

$$\mathfrak{L} = \nabla_M^2 - \frac{m^2}{r^2} + \frac{2}{r} \bar{u}_r \operatorname{div}_M$$

Scalar product (29), and the metric derived from it, define a Hilbert space. The main properties of transformation $\mathfrak{L}$ are obtained from a consideration of $\langle \bar{v}, \mathfrak{L}\bar{v} \rangle$, where $\bar{v}$ belongs to the domain of vectors satisfying the boundary conditions appearing in (29). If we apply (31) to $\bar{v}$, we discover that the right-hand member vanishes, so that

$$\langle \bar{v}, \mathfrak{L}\bar{v} \rangle = \iint_D \left[\bar{v} \cdot \mathfrak{L}\bar{v} + \frac{r^2}{m^2} \operatorname{div}_M \bar{v} \cdot \operatorname{div}_M \mathfrak{L}\bar{v} \right] r dr dz$$

$$= -\iint_D \left\{ \left| \operatorname{curl}_M \bar{v} \right|^2 + \left| \frac{m}{r} \bar{u}_\phi \times \bar{v} \right. \right.$$

$$\left. \left. + \operatorname{curl}_M \left(\frac{r}{m} \bar{u}_\phi \operatorname{div}_M \bar{c} \right) \right|^2 \right\} r dr dz. \tag{50}$$

Clearly, $\langle \bar{v}, \mathfrak{L}\bar{v} \rangle$ is never positive. We now want to prove that $\mathfrak{L}\bar{v}_0 = 0$ implies $\bar{v}_0 = 0$, which would then make transformation (28) negative-definite. We first need to establish Helmholtz' theorem in the meridian plane. More explicitly, we want to examine the splitting of a meridian vector $\overline{P}$ into

$$\overline{P} = \operatorname{grad} A + ; \tag{51}$$

where grad A, the longitudinal term, is required to be perpendicular to (c), and to have the same divergence as $\overline{P}$. In other words, A must satisfy

$$\nabla_M^2 A = \operatorname{div}_M \overline{P} \qquad A = 0 \text{ on } (C).$$

It is a simple matter, with the help of Green's theorem, to show that this problem has a unique solution, and that the longitudinal term vanishes when $\operatorname{div}_M \overline{P} = 0$. It is also a simple matter, using Stokes' theorem in the meridian plane, to show that each meridian vector for which curl $\overline{P} = 0$ can be put in the form $\operatorname{grad}_M \theta$. If, in addition, $\overline{P}$ is perpendicular to the boundary c, potential θ is nothing but the function A appearing in (51). The sources of $\bar{t}$ are, consequently, the curl of $\overline{P}$ and the tangential components of $\overline{P}$. When $\mathfrak{L}\bar{v}_0 = 0$, the left-hand member of (51) vanishes; this implies that the squares in the second member also vanish, and, in particular, that curl $\bar{v}_0 = 0$. Letting $\bar{v}_0 = \operatorname{grad} A$, it is found that A must satisfy

$$\mathfrak{L}\bar{v}_0 = \nabla_M^2 \operatorname{grad} A - \frac{m^2}{r^2} \operatorname{grad} A + \frac{2}{r} \bar{u}_r \nabla_M^2 A = 0$$

with

$$A = \nabla_M^2 A = 0 \text{ on } (C).$$

Projection of $\mathfrak{L}\bar{v}_0$ on the z axis indicates that

$$\frac{\partial}{\partial z}\left[\nabla_M^2 A - \frac{m^2}{r^2} A\right] = 0.$$

In consequence, $\nabla_M{}^2 A - (m^2/r^2)A$ has a constant value along a parallel to the z axis. This value must be zero, because

$$\nabla_M^2 A - \frac{m^2}{r^2} A = 0 \tag{52}$$

on c. It follows that (52) is valid over the whole area D. An application of Green's theorem shows that

$$\iint_D \left\{\frac{m^2}{r^2} A^2 + |\operatorname{grad} A|^2\right\} r dr dz = 0$$

so that both A and $\bar{v}_0$ must vanish.

The self-adjoint character of $\mathfrak{L}$ (*i.e.*, $\langle \bar{c}, \mathfrak{L}\bar{d}\rangle = \langle \bar{d}, \mathfrak{L}\bar{c}\rangle$) can be quickly established by using a relation derived from the three-dimensional Green's theorem (32):

$$\iint_D \left\{\overline{Q}\cdot\left(\nabla_M^2 \overline{P} - \frac{m^2}{r^2}\overline{P} + \frac{2mg}{r^2}\bar{u}_r\right) + h\left(\nabla_M^2 g - \frac{m^2+1}{r^2} g + \frac{2mc_r}{r^2}\right)\right.$$
$$+ \left(\operatorname{div}_M \overline{P} - \frac{mg}{r}\right)\cdot\left(\operatorname{div}_M \overline{Q} - \frac{mh}{r}\right) + \operatorname{curl}_M \overline{P}\cdot\operatorname{curl}_M \overline{Q}$$
$$\left. + \left[\operatorname{curl}_M (g\bar{u}_\phi) + \frac{m}{r}(\bar{u}_\phi \times \overline{P})\right]\cdot\left[\operatorname{curl}_M (h\bar{u}_\phi) + \frac{m}{r}(\bar{u}_\phi \times \overline{Q})\right]\right\} r dr dz$$
$$= \int_c \left\{(\bar{u}_n \times \overline{Q})\cdot\operatorname{curl}_M \overline{P} - \frac{hm}{r}(\bar{u}_n\cdot\overline{P}) - h\bar{u}_t\cdot\operatorname{curl}_M (g\bar{u}_\phi) + (\bar{u}_n\cdot\overline{Q})\left(\operatorname{div}_M \overline{P} - \frac{mg}{r}\right)\right\} r dc. \tag{53}$$

If we use this relation twice, setting $g = r/m \operatorname{div}_M \overline{P}$, $h = r/m \operatorname{div}_M \overline{Q}$ and subtracting, we obtain, since

$$\operatorname{div}_M\left[\nabla_M^2 \bar{c} - \frac{m^2\bar{c}}{r^2} + \frac{2}{r}\cdot\bar{u}_r \operatorname{div}_M \bar{c}\right]$$
$$= \frac{m}{r}\nabla_M^2\left(\frac{r \operatorname{div}_M \bar{c}}{m}\right) + \frac{2m^2}{r^3} c_r - \frac{m^2+1}{r^2}\operatorname{div}_M \bar{c}$$
$$= \nabla_M^2 (\operatorname{div}_M \bar{c}) - \frac{m^2}{r^2}\operatorname{div}_M \bar{c} + \frac{2m^2}{r^3} c_r + \frac{2}{r}\frac{\partial}{\partial r}(\operatorname{div}_M \bar{c}), \tag{54}$$

the relation,

$$\iint_D \left[\overline{Q}\cdot\mathfrak{L}\overline{P} - \overline{P}\cdot\mathfrak{L}\overline{Q} + \frac{r^2}{m^2}\operatorname{div}_M \overline{Q} \operatorname{div}_M \mathfrak{L}\overline{P}\right.$$
$$\left. - \frac{r^2}{m^2}\operatorname{div}_M \overline{P} \operatorname{div}_M \mathfrak{L}\overline{Q}\right] r dr dz = 0. \tag{55}$$

The second member in (53) vanishes because of the boundary conditions. Eq. (55) is nothing but $\langle \bar{c}, \mathfrak{L}\bar{d}\rangle - \langle \bar{d}, \mathfrak{L}\bar{c}\rangle = 0$, the relation we set out to prove.

A Printed Circuit Balun for Use with Spiral Antennas*

R. BAWER† AND J. J. WOLFE†

Summary—**A novel printed circuit balun is described which is particularly well suited to applications where space is at a premium. The design utilizes unshielded strip transmission line, but is readily adaptable to all of the common printed circuit transmission line techniques. When the balun is housed within the cavity of a spiral antenna, boresight error is virtually eliminated, ellipticity ratios of less than 2 db are maintained over an azimuth angle greater than ±60°, and the input standing-wave ratio is less than 2:1 over an octave frequency range. Experimental results are given and additional applications are described.**

* Manuscript received by the PGMTT, October 16, 1959; revised manuscript received, November 23, 1959. The work reported in this paper was sponsored by the Airborne Instruments Lab., Div. of Cutler-Hammer, Inc., Melville, N. Y., P.O. No. 6468, under Air Force Contract No. AF33(600)-37929.

† Aero Geo Astro Corp., Alexandria, Va.

I. Introduction

A BALUN is a term used by antenna engineers to describe a device which transforms an unbalanced to a balanced transmission line. To the microwave engineer, the same device might be called a ratrace, magic tee, or more generally a hybrid. In lumped circuit applications, we also find a similar device used in conjunction with balanced mixers, phase detectors, and single-sideband modulators, to name a few. However, regardless of what the device is called, the operation will appear to be basically similar provided the analysis is made using a compatible frame of reference. One particularly powerful tool used at microwaves

involves the circuit symmetry; however, for the particular baluns to be described, this approach is somewhat mathematical in nature and tends to obscure the physical significance of the various parameters. On the other hand, the equivalent circuit representation clearly places in evidence the role of each element of the unit and provides the designer with a compact analytic expression for the input impedance.

There are numerous devices[1–5] which are suitable for transforming a balanced to an unbalanced transmission line over a wide frequency range. The balun to be described here is nothing more than a novel realization of a conventional type of balun which is particularly suited for use with spiral antennas or similar applications where space is at a premium.

Fig. 1 shows the schematic drawing and equivalent circuit of a coaxial balun proposed by Roberts.[1] Z_a and Z_b represent the characteristic impedance of the transmission lines a and b, respectively. Z_{ab} is the characteristic impedance of the balanced transmission line composed of the outer conductors of lines a and b.

From the equivalent circuit, the impedance looking into the terminals D–G can be directly written as

$$Z = -jZ_b \cot \theta_b + \frac{jRZ_{ab} \tan \theta_{ab}}{R + jZ_{ab} \tan \theta_{ab}} \cdot \tag{1}$$

On letting $\theta_b = \theta_{ab} = \theta$ and rationalizing, the following expression is obtained

$$Z = \frac{RZ_{ab}^2 + j \cot \theta [R^2(Z_{ab} - Z_b \cot^2 \theta) - Z_b Z_{ab}^2]}{Z_{ab}^2 + R^2 \cot^2 \theta} \cdot \tag{2}$$

Two interesting cases have been treated in the literature. Roberts[1] imposes the conditions that $Z_b = Z_a$ and $Z_{ab} = R$. For this case,

$$Z = R \sin^2 \theta + j(\cot \theta)(R \sin^2 \theta - Z_a) \tag{3}$$

and the input impedance becomes perfectly matched at two widely-separated frequencies given by the solution of

$$\sin^2 \theta = Z_a/R. \tag{4}$$

These frequencies are symmetrically disposed about a center frequency corresponding to $\theta = 90°$.

Using essentially the same prototype balun, McLaughlin, *et al.*,[2] obtained a considerably greater bandwidth by minimizing the imaginary term in (2). Again, θ is chosen to be 90° at midband, Z_b is chosen according to

$$Z_b = R^2/Z_{ab}, \tag{5}$$

and Z_{ab} is made as large as possible.

In order to compare the two approaches, the standing-wave ratio curves shown in Fig. 2 were calculated for the case of a balun transforming a 50-ohm unbalanced load to a 70-ohm balanced load. The improvement of the more flexible form (*i.e.*, $Z_b = R^2/Z_{ab}$) is obvious. The advantage is somewhat offset by the increased complexity of the coaxial realization; however, this objec-

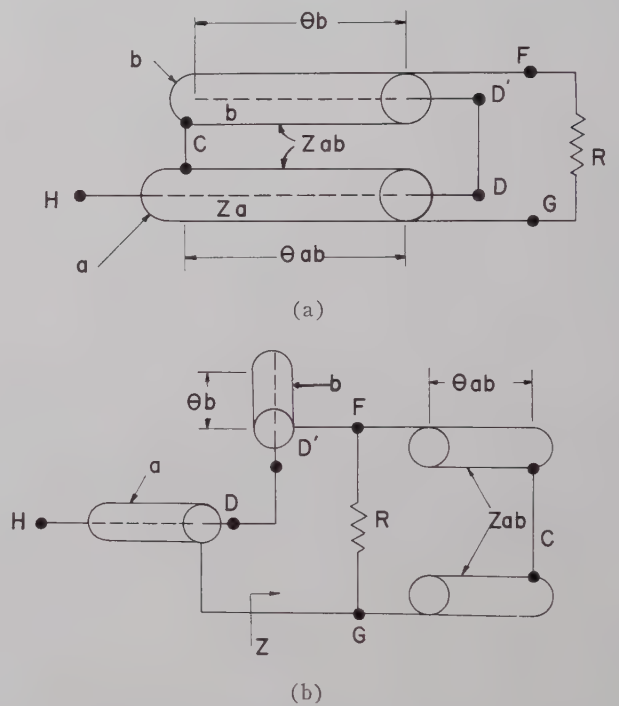


Fig. 1—Wide-band balun, from Roberts.[1] (a) Schematic drawing. (b) Equivalent circuit.

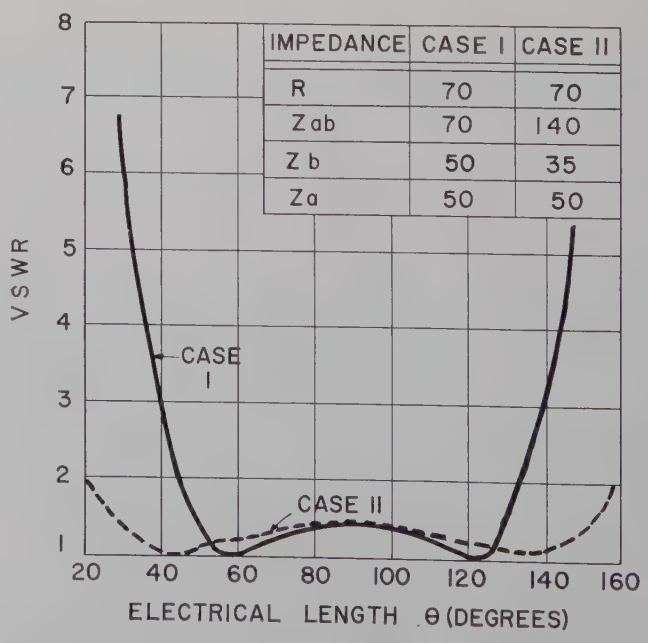


Fig. 2—Theoretical VSWR vs electrical length for two different balun design criteria.

[1] W. K. Roberts, "A new wide-band balun," PROC. IRE, vol. 45, pp. 1628–1631; December, 1957.
[2] J. W. McLaughlin, D. A. Dunn, and R. W. Grow, "A wide-band balun," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-6, pp. 314–316; July, 1958.
[3] Radio Res. Lab. Staff, "Very High Frequency Techniques," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 85–91; 1957.
[4] E. G. Fubini and P. J. Sutro, "A wide-band transformer from an unbalanced to a balanced line," PROC. IRE, vol. 35, pp. 1153–1155; October, 1947.
[5] E. M. T. Jones and J. K. Shimizu, "A wide-band strip-line balun," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-7, pp. 128–134; January, 1959.

tion is no longer valid for the printed circuit balun.

It is important to note that, in both of the cases illustrated, the balun provides no impedance transformation at the center frequency. This fact is particularly significant in all applications in which the load impedance is considerably different than 50 ohms. A low standing-wave ratio can then be obtained only by the use of an independent impedance transformer.

The need for an impedance transformer in a small diameter coaxial line again suggests the use of printed circuit techniques; in this case, to alleviate the severe mechanical problems and high cost which might otherwise be incurred. Still another incentive stems from an assumption inherent in the equivalent circuit representation of the balun. In this analysis, it was tacitly assumed that the line length D–D' is small compared to the operating wavelength. The simultaneous requirements that Z_{ab} be large and the length D–D' be small, are obviously in conflict; moreover, this length becomes prohibitively long at frequencies greater than approximately 2000 mc.

To circumvent this difficulty, one can introduce right-angle bends in the coaxial lines a and b, and thereby reduce length D–D' to an arbitrarily small value. It is this coaxial structure that served as the prototype for the printed circuit balun.

II. Printed Circuit Balun

Fig. 3 illustrates the construction of the printed circuit balun. Note that the coaxial line of Fig. 1 has been replaced by unshielded strip transmission line and that the corresponding balanced transmission line now consists of the two ground planes.

For a given load and generator impedance, (2), (4), or (5) can be used to calculate the necessary parameter values. If, however, the mismatch corresponding to the ratio of load-to-generator impedance is greater than can be tolerated, an independent transforming section must be used. This can easily be accomplished by tapering the input transmission line.

The remaining steps of the design involve translating the calculated impedances and line lengths into actual physical dimensions of the printed line. At this point, it is important to consider any and all constraints which must be imposed on the balun. These constraints may be clearly defined, or may be contained within a broader specification, as in the case of the antenna assembly described in the next section.

With reference to Fig. 3, it is seen that two types of transmission line are involved. The first is the conventional unshielded strip transmission line consisting of a thin conductor over a ground plane; namely, lines a and b. The second type consists of a balanced transmission line made up of two flat, thin conductors of characteristic impedance Z_{ab} and length θ_{ab}. The simpler case will be discussed first.

The characteristic impedance and phase velocity of unshielded strip transmission line has been the subject of numerous articles.[6,7] The important point to remember is that virtually all of the field is confined within a region of about three conductor widths. Consequently, if the ground plane width is taken equal to or greater than three times the conductor width ($B>3b$), the resulting transmission line is, for all practical purposes, nonradiating. As this criterion is violated (say $b<B<3b$), the configuration takes on the appearance of a balanced transmission line and care must be exercised to minimize radiation caused by the unbalanced-to-balanced junction.

For reference purposes, approximate dimensions of unshielded strip transmission line for typical characteristic impedances are given in Table I. The reader is directed to the references for more complete data.

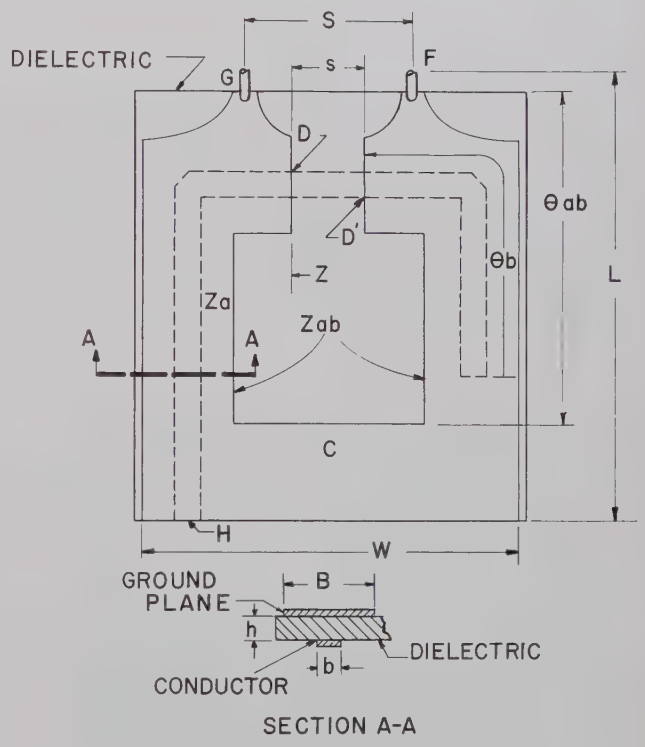


Fig. 3—Illustration of the construction of a printed circuit balun.

TABLE I

Characteristic Impedance of Unshielded Strip Transmission Line

($W>b$, $t=0$)

b/h	Material G-10 Z_0 (ohms)	Material GB-112 Z_0 (ohms)
0.75	90	125
1.0	80	112
1.5	65	89
2.0	53	74
2.5	46	64
3.0	41	55
4.0	33	46

[6] M. Arditi, "Characteristics of microstrip for microwave wiring," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 31–56; March, 1955.

[7] F. Assadourian and E. Rimai, "Simplified theory of microstrip transmission systems," Proc. IRE, vol. 40, pp. 1651–1657; December, 1952.

The physical length of the open-circuited transmission line (θ_b) depends upon the phase velocity in the printed circuit line. For the case of a homogeneous dielectric surrounding the lines, this phase velocity is given by

$$v = c/\sqrt{\epsilon} \tag{6}$$

where c is the velocity of light in free space and ϵ is the relative dielectric constant of the medium. Since the region between the conductor and the ground plane is filled with a dielectric while the remaining cross section is air, the problem becomes quite complex and one generally resorts to empirical techniques to establish the phase velocity. The velocity decreases as the width of the conductor increases and asymptotically approaches the velocity which would be obtained if the line were completely immersed in the dielectric.

In practice, the range of the ratio of conductor width to spacing above the ground plane is $1 < (b/h) < 5$. Over this range the phase velocity varies by about 10 per cent, which is well within the accuracy required for the balun design. Therefore, one may elect to use an average phase velocity (or average effective dielectric constant) instead of making the series of measurements necessary to describe the actual behavior. To this end, an approximate average effective dielectric constant can be taken as

$$\epsilon_e = 0.75\epsilon, \tag{7}$$

where ϵ is the dielectric constant of the base material. Eq. (6) can then be used for the initial calculation of θ_b, and the final length may easily be obtained by judicious trimming of the line length.

The impedance Z_{ab} and its corresponding phase velocity θ_{ab} present a more difficult problem. Short of actually attempting a solution to the problem using conformal mapping techniques, there appears to be no simple way to calculate or estimate the impedance of this configuration using available data. To further compound the problem, it may be recalled that the derivation of (1) was based on the assumption of a uniform balanced transmission line. The correct expression for the impedance Z of Fig. 3 must take into account not only the change in conductor spacing, but also the shape, length, and spacing of the terminals G–F. This is clearly no small undertaking.

In short, an analytic or approximate expression for the value of Z_{ab} is very desirable. Fortunately, the absence of this information has not proved detrimental, as attested to by the experimental results discussed in the next section. The redeeming feature can be attributed to the design requirement that Z_{ab} be as large as possible.

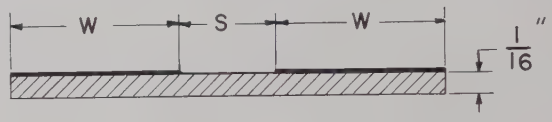
Nevertheless, some idea of the order of magnitude of Z_{ab} would be useful. To this end, measurements were taken over the range of strip widths and spacings which might be encountered in practice. The data were taken with a Boonton type 190 Q meter, using conventional lumped element measurement techniques to obtain the inductance and capacitance of each configuration. The characteristic impedance was then calculated from the relation

$$Z_{ab} = \sqrt{L/C} \text{ ohms}. \tag{8}$$

The phase velocity does not significantly depart from the free space value, a fact which is to be expected in light of the line configuration. The results are given in Table II and should be sufficiently accurate for most balun designs.

TABLE II

Characteristic Impedance of Printed Balanced Transmission Line



W/S	Z_0 (ohms)
2	220
4	185
6	170
8	160
10	155

III. Application to Spiral Antennas

The balun described here was developed to meet the need for a wide-band, compact device to feed a dual-arm spiral antenna from an unbalanced coaxial transmission line. It should be noted in passing that the printed circuit realization is not limited to this particular application, but is readily adaptable to a wide variety of uses. The printed circuit balun for spiral antennas represents one of the most stringent applications of this technique and, as such, clearly illustrates its extreme versatility.

At the onset of the development program, it was decided that the balun should be housed within the antenna structure. This necessarily imposed certain constraints on the physical size of the balun. In order to obtain a better insight into these constraints, a brief review of the operation of spiral antennas is given. The discussion is limited to those aspects which affect the balun design, and vice versa.

The spiral antenna[8] is one of a class of so-called "frequency independent" antennas which are completely described by angles except for diameter. Physically, the antenna is a two-dimensional structure which is fabricated by photo-etching the geometric configuration on copper-clad laminate. Of the wide variety of shapes possible, only three have received wide attention: the

[8] The spiral antenna was developed by W. E. Turner at Wright Air Dev. Center, Dayton, Ohio, in 1954. Since that time, most of the spiral antenna discussions have appeared as classified documents. A good unclassified reference is J. A. Kaiser, "Spiral antennas applied to scanning arrays," *Electronic Scanning Symp.*, Air Force Cambridge Res. Center, Bedford, Mass.; April, 1958. See also R. Bawer and J. J. Wolfe, "The Spiral Antenna," presented at 1960 IRE International Convention, New York, N. Y.; March 21–24, 1960. To be published in the 1960 IRE International Convention Record.

logarithmic or equiangular spiral, the Archimedean spiral, and the rectangular counterpart of the Archimedean spiral. The printed antenna may be of the single or two-conductor configuration, the latter being preferable at frequencies below about 5000 mc. Functionally, the antenna radiates a bidirectional circularly polarized beam normal to the plane of the printed element when fed from a balanced two-wire transmission line. In most applications, the spiral antenna is flush-mounted (*e.g.*, in the skin of an aircraft), so that the bidirectional characteristics are undesirable; unidirectional patterns may be obtained by mounting the spiral at the mouth of a cylindrical cavity as shown in Fig. 4.

The gain of the cavity-backed spiral antenna follows the behavior one would expect from a dipole over a ground plane, being a maximum when the cavity depth is about $\frac{1}{4}\lambda$ and dropping off quite rapidly for depths less than the $\frac{1}{8}\lambda$ and greater than $\frac{3}{8}\lambda$. Since it is desirable to house the balun within the cavity, the length of the structure must be less than one-eighth of the free-space wavelength for maximum bandwidth. This requirement is based on the assumption that the balun will be mounted at right angles to the plane of the printed spiral as shown in the figure.

The allowable width of the balun is related to the radiating mechanism of the antenna. Although there is no rigorous theory to explain the radiation from spirals, the "Band Theory" quite adequately describes the phenomenon and is generally accepted by workers in the field. According to this theory, radiation from a tightly-wound spiral will occur from a band with mean diameter equal to λ/π. Therefore, the width of the balun structure in the vicinity of the spiral terminals should be considerably less than this diameter. If this criterion is violated, obscuration of the spiral occurs and the axial ratio of the antenna is severely affected. In fact, axial ratio measurements at the higher frequencies afford an excellent means for checking this criterion. Experimental data have indicated that satisfactory operation will result if the width of the balun in the vicinity of the antenna terminals is less than about $\lambda/8$.

In review, there are constraints which are imposed on the balun size by virtue of its use with a spiral antenna and the manner in which the balun is mounted in the antenna cavity. For the configuration shown in Fig. 4, the criteria $L \leq \frac{1}{8}(\lambda)$ maximum and $W \leq \frac{1}{8}(\lambda)$ minimum prove to be a useful rule of thumb. The dimension S, the spacing between the output terminals of Fig. 3, is made as small as practical and compatible with the input terminal spacing of the spiral antenna.

Finally, it seems reasonable to evaluate any balun in terms of its behavior under actual operating conditions; namely, when driven from a prescribed generator and terminated in prescribed load. Thus, the final evaluation of the printed balun feed for the spiral antenna should be based on the characteristics of the entire antenna structure. Typical specification for spiral antennas of the type considered are:

Frequency range—1 octave,
Axial ratio—3 db maximum,
VSWR—2:1 maximum,
Beam squint—$<5°$.

In the following section, some experimental results are discussed. The data are limited to the antenna configuration shown in Fig. 4 in which the balun is contained within the cavity and mounted at right angles to the plane of the spiral. Several alternate balun configurations and additional applications of the printed circuit balun are given in Section V.

IV. Experimental Results

A series of printed circuit baluns were developed to operate with cavity-backed spiral antennas over the frequency range of 104 mc through 4200 mc. The initial design effort was concentrated on the *S*-band unit, since this case represented the most severe situation from the point of size and, consequently, the most severe test of the validity of the approximations.

Fig. 5 illustrates the final form of the balun feed which was designed to operate with a 3-inch diameter, dual-arm, Archimedean spiral antenna. The balun was photo-etched on 1/32 inch thick, 2-ounce copper-clad epoxy fiberglass (G-10) and is approximately 1 inch long by $\frac{3}{4}$ inch wide.

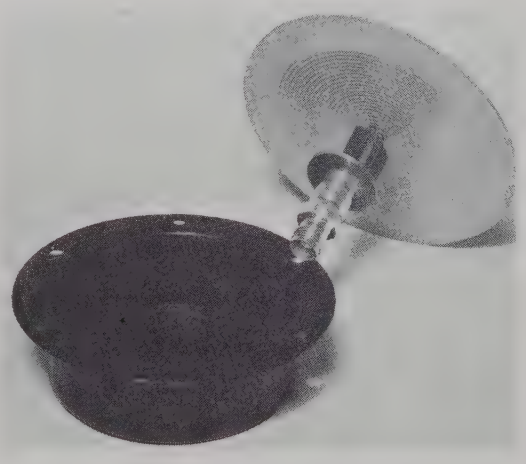


Fig. 4—An exploded view of a spiral antenna assembly.

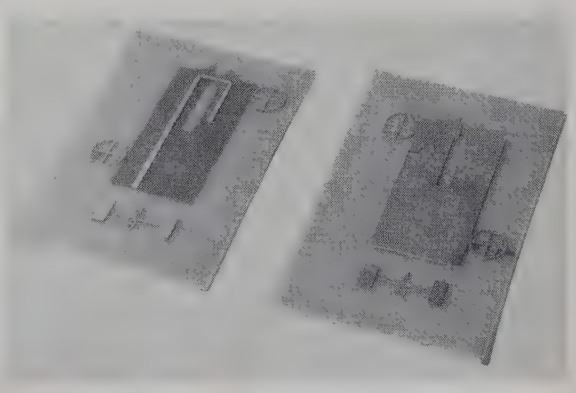


Fig. 5—Photograph of the final balun feed for an *S*-band spiral antenna.

Attention is specifically called to the following points:

1) The impedance of the spiral antenna was estimated to be in excess of 100 ohms; consequently, the input conductor was tapered to provide a characteristic impedance of about 50 ohms at the balun input and an impedance of about 100 ohms at the center line of the balun. The line continues at this impedance to the open circuit.
2) At the upper end of the frequency band, obscuration of the spiral by the balun resulted in an increased axial ratio. This effect was virtually eliminated by the simple expedient of trimming the corners of the ground plane as shown in Fig. 4.
3) The final results were obtained by systematically varying a number of the parameters. The major change involved the length of the open-circuited line θ_b which was adjusted to give a VSWR curve that is roughly symmetrical about the center frequency. The characteristic impedance and length of the balanced line Z_{ab} and θ_{ab} were found to have only a second-order effect.

Data showing the frequency dependence of the input impedance and axial ratio of a printed circuit balun and an *S*-band spiral antenna are shown in Fig. 6. If we neglect antenna gain, the major factor which limits the band-width of the antenna assembly is the axial ratio. The deterioration of axial ratio can be attributed to the spiral diameter at the low end of the frequency band (not illustrated) and balun size at the high end of the band. The input impedance is, of course, adversely affected at these limits also, but to a lesser extent than the axial ratio.

Fig. 7 illustrates the type of radiation patterns which have been obtained. The data were taken for two orthogonal cuts on an automatic pattern recorder with the antenna immersed in microwave absorbing material to simulate free space operating conditions. Axial ratio data (Fig. 6) were obtained by recording the maximum variation in the power received from a test antenna located on axis ($\phi = \theta = 0$ degrees) as the spiral antenna was rotated 360°; this corresponds to the maximum difference (in db) between E_θ and E_ϕ at the origin of Fig. 7 for all orientations of the spiral.

There are two important points which should be emphasized. The first relates to range of azimuth angles over which the ellipticity ratio (E_0/E_θ) remains comparable to the on-axis ratio. For the particular cases illustrated, the ellipticity ratio remains less than 1 db over an azimuth angle of greater than $\pm 60°$.

The second point relates to the symmetry of the patterns about the axis of the antenna; beam squint is negligible. This behavior should be compared to that obtained with more conventional feeds in which beam squint in the order of 10° to 15° is common.

The results of the *S*-band balun are typical of the data which have been taken on units designed to operate as low as 140 mc. The chief difference between these baluns and the *S*-band unit is the thickness of the copper-clad material; for mechanical reasons, the lower-frequency baluns were photo-etched on 1/16-inch material.

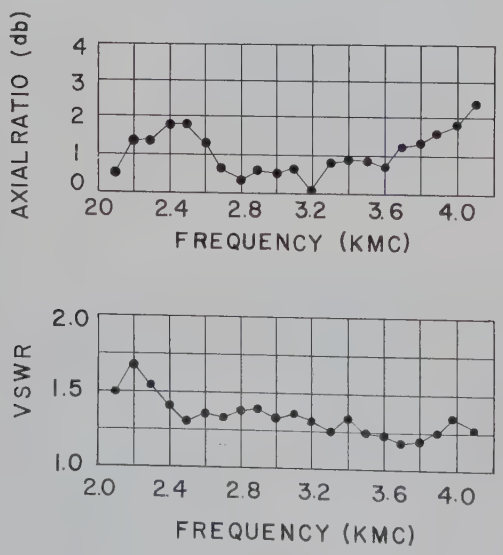


Fig. 6—Input impedance and axial ratio of a printed circuit balun and spiral antenna.

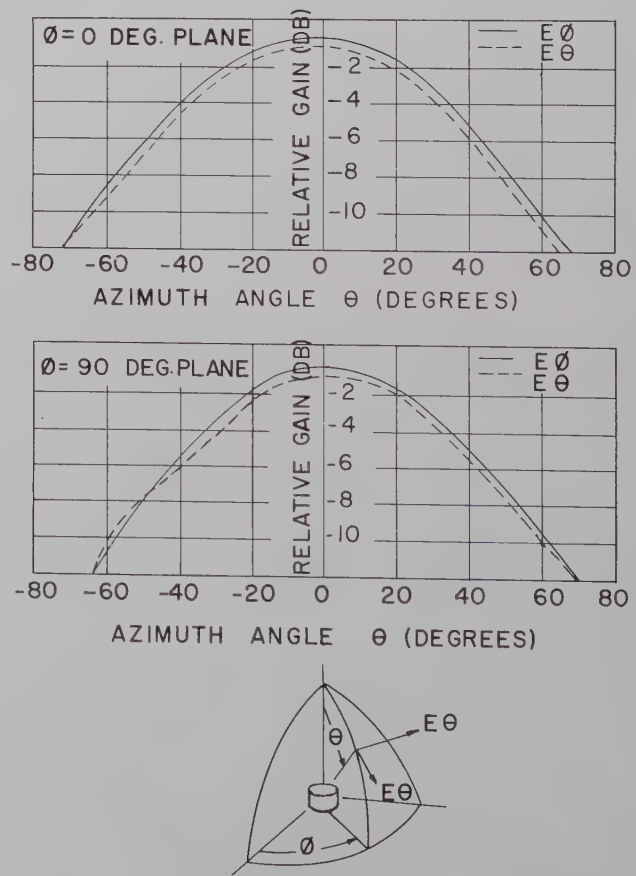


Fig. 7—Typical radiation patterns of an *S*-band spiral antenna fed from a printed circuit balun.

V. Conclusions

In the preceding discussion, a rather specialized application of a printed circuit balun was presented. The use of the balun with a cavity-backed spiral antenna, particularly when the balun is housed within the cavity, places severe limitations on the element size; it is expected that improved performance might be realized if these constraints were relaxed.

Nevertheless, this form factor does have some definite advantages which might not be immediately apparent. For example, by simply joining a second printed line at right angles to the balun (*i.e.*, normal to the printed line), the entire back plate of the cavity is made available for additional circuitry. In this manner, directional couplers, filters, etc., can easily be incorporated directly within the antenna structure with no space and negligible weight penalty.

An alternate configuration which appears to have some merit would involve mounting the balun parallel to the spiral and close to the cavity back plate. For this case, the size constraints can be significantly relaxed and the impedance transformer could now be placed in a printed balanced line connecting the balun output to the spiral input terminals. In other applications, it may be desirable to utilize a shielded configuration. The design described is equally applicable to all of the shielded strip transmission lines in common use.

In conclusion, it should be emphasized that the printed circuit balun is inherently capable of performance comparable to that of other wide-band baluns. With regard to basic design, nothing new is claimed. With regard to physical realization, it is believed that this printed circuit technique offers much in the way of ease of fabrication and miniaturization and, above all, a flexibility found in no other transmission line.

The *P-I-N* Modulator, an Electrically Controlled Attenuator for MM and Sub-MM Waves*

F. C. DE RONDE†, H. J. G. MEYER†, and O. W. MEMELINK†

Summary—**The construction and performance of a millimeter wave modulator are described. The main part of the modulator consists of a *p-i-n* germanium structure inserted into a rectangular waveguide. A modulation depth of 11 db could be obtained at frequencies up to 5 kc, this modulation being caused for the greatest part by attenuation.**

I. Introduction

MICROWAVE modulators are used for various purposes, *e.g.* the formation of sidebands and particularly the increase of the sensitivity of measurements. In view of the great importance of having a modulator available, it is not surprising that various types of modulators are already in use to date.

The one most commonly employed is an absorption type modulator which makes use of a ferrite. Here the losses of the ferrite or the plane of its polarizing action are varied by a magnetic field. The necessity of using a magnetic field limits the modulating frequency to rather low values and the coils required make the device somewhat bulky.

Another type of modulator, also employing a magnetic field, was proposed by Gunn and Hogarth.[1] Here the number of free charge carriers in a semiconductor is varied by driving them either to a surface with a high or to a surface with a low value of the surface recombination velocity.

A modulator in which no use is made of a magnetic field is the so-called transparitor of Arthur, Gibson, and Granville.[2] Here the differential mobility of charge carriers in a semiconductor, and therefore its attenuation, is varied by the application of a strong electric field. An advantage of this arrangement is the possibility of employing very high modulation frequencies, the theoretical response time of the device being of the order of 10^{-12} seconds. A definite disadvantage for use under continuous wave conditions, however, is its great

* Manuscript received by the PGMTT, November 30, 1959; revised manuscript received, January 18, 1960.

† Philips Res. Labs., N. V. Philips' Gloeilampenfabrieken, Eindhoven, Netherlands.

[1] J. B. Gunn and C. A. Hogarth, "A novel microwave attenuator using germanium," *J. Appl. Phys.*, vol. 26, p. 353; March, 1955.

[2] J. B. Arthur, A. F. Gibson and J. W. Granville, "The effect of high electric fields on the absorption of germanium at microwave frequencies," *J. Electronics*, vol. 2, p. 145; September, 1956.

power consumption, which gives rise to cooling problems.

Finally, very useful modulators can be made from various types of semiconductor point-contact diodes.[3–5] As in this case the modulation of the microwave signal is caused mainly by changing the reflectivity of the device, this type of modulator has to be used in combination with an isolator or circulator. If such devices are available, point-contact diode modulators show none of the disadvantages mentioned above, as they are small devices with low power consumption allowing very high modulation frequencies.

The *p-i-n* modulator to be described below[6,7] has advantages similar to those of the point-contact diode modulators except that the frequency of modulation is limited to rather low values. On the other hand, it has the advantage that it can be used without an isolator or circulator because its action is based mainly on a varying attenuation rather than on a varying reflection. This implies that it can also be used with advantage as a variable attenuator in control systems. The principle on which the action of *p-i-n* modulator is based is described below.

II. Design Considerations

If the intrinsic region of a semiconductor *p-i-n* structure is placed into a waveguide while care is taken that the heavily doped *p* and *n* regions are well outside the guide, the microwave field undergoes only the influence of a nearly insulating dielectric. By applying a forward voltage to contacts on the *n* and *p* regions, electrons and holes are injected into the intrinsic (or at least practically intrinsic) region and microwave power is absorbed by the mobile charge carriers thus introduced into the intrinsic part. The mechanism of transport of charge carriers into this region is governed by diffusion and recombination. Theory shows that the excess-hole and electron concentrations are almost homogeneous if the distance between the *p* and *n* regions is not larger than $L=\sqrt{D\tau}$, the diffusion-recombination length. Here D is the ambipolar diffusion constant and τ the average life time of electron-hole pairs. The time necessary to establish a certain concentration pattern is also of the order of τ. Hence τ determines the maximum modulation-frequency obtainable. From these considerations conclusions can be drawn regarding the width of the *i* region which is to be used in the modulator structure. In fact, as it is desirable to get an appreciable variation of the charge carrier density over all the material within the waveguide, the width of the intrinsic region between the *p* and *n* regions should not be much larger than the recombination-diffusion length L.

Using for germanium the values $D=60$ cm²/sec and $\tau\approx 5$ μsec we find $L\approx 0.2$ mm. This small diffusion length implies that the distance of the broad faces of the waveguide should be of the same order of magnitude, from which it is clear that the most obvious application of this type of modulator is in the millimeter and submillimeter wave range.

III. Construction of a *P-I-N* Modulator for Use at 4 mm

Into a wafer of germanium, a gold and an aluminum wire are alloyed parallel to each other as shown in Fig. 1. The germanium, which has an impurity content of about 10^{12} cm^{-3}, is intrinsic at room temperature and has a resistivity of about 50 ohm-cm. The gold wire contains a small amount of arsenic and, as arsenic is a donor and aluminum an acceptor in germanium, the recrystallized material underneath the alloyed wires provides heavily doped *n*- and *p*-type regions.

Following the alloy process the wafer is etched in a conventional manner. After washing and drying, nickel wires of 50-μ diameter are attached to the gold and aluminum contacts. Finally the wafer is imbedded in a mylar foil in order to facilitate the mounting of the *p-i-n* structure.

The modulator is inserted into the 4-mm waveguide through slots in the broad faces in such a way that only the intrinsic region is inside the waveguide (see Fig. 2). Because this region is rather small the waveguide has been tapered. Use can be made also of tapered ridged waveguides or striplines. In order to adjust the intrinsic region just inside the waveguide, a simple construction was made to move the *p-i-n* structure up and down.

IV. Performance

An absorption variation of about 11 db was obtained with a dc current variation of about 15 ma and a maximum power consumption of about 10 mw. The insertion losses were a few decibels, due to the rather low resistivity (50 ohm-cm) of intrinsic germanium.

As can be seen from Fig. 3, at modulation frequencies up to about 5 kc the response is practically independent of frequency while the modulation depth has decreased by 3 db at 50 kc. From this the life time τ may be estimated to be of the order of 3 μsec.

Since the length of the active region in the waveguide, 2.5 mm, was about half a wavelength and, at the cost of higher insertion losses, can be made even longer, the response of the modulator covers the whole fre-

[3] M. A. Armistead, E. G. Spencer, and R. D. Hatcher, "Microwave semiconductor switch," Proc. IRE, vol. 44, p. 1875; December, 1956.

[4] R. V. Garver, E. G. Spencer, and R. C. LeCraw, "High-speed microwave switching of semiconductors," *J. Appl. Phys.* vol. 28, p. 1336; November, 1957.

[5] K. J. S. Cave, W. Neu, and A. C. Sim, "A diode modulator for mm-waves," *Proc. IEE*, vol. 106, 1959. (Communication given at the International Convention on Transistors and Associated Semiconductor Devices, London, Eng.; May 21–27, 1959.)

[6] Dutch patent application no. 229531, 11.7.1958.

[7] A. Uhlir, Jr., "The potential of semiconductor diodes in high-frequency communications," Proc. IRE, vol. 46, pp. 1099–1115; June, 1958.

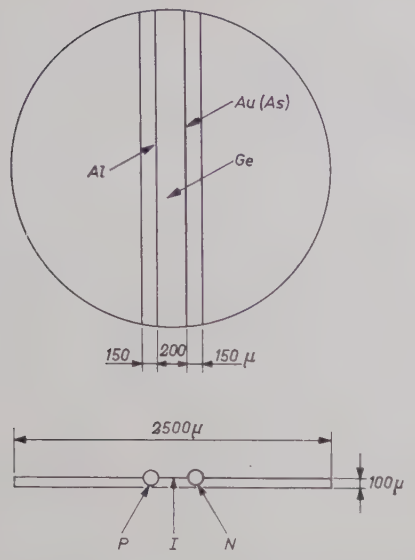


Fig. 1—Geometry of the *p-i-n* germanium wafer.

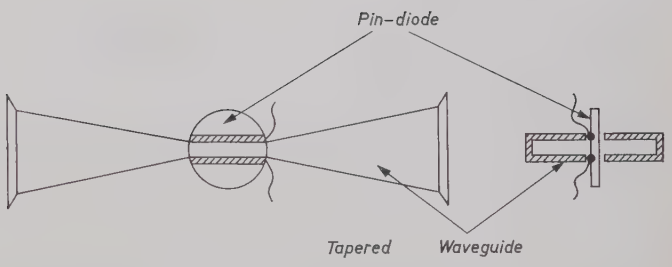


Fig. 2—Position of the modulator in the waveguide.

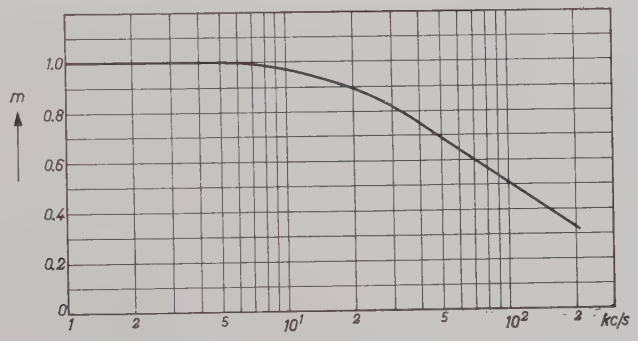


Fig. 3—Modulation depth vs modulation frequency.

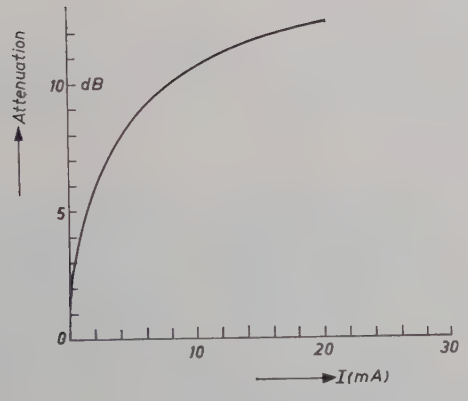


Fig. 4—Attenuation vs modulator current.

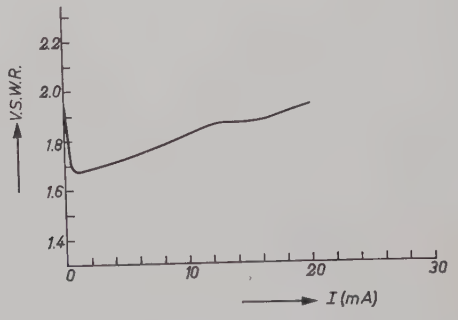


Fig. 5—Voltage standing wave ratio.

quency band of the klystron (67–77 kmc/sec); in fact the device was also successfully used at 120 kmc/sec.

In Fig. 4 the attenuation is given as a function of the current through the modulator. Because the intrinsic region is not tapered there is some reflection. However, as can be seen from Fig. 5, this reflection, measured with the modulator being terminated by a matched load, is nearly independent of the modulator current. A theoretical analysis of the relation between the modulator current and the attenuation was not made because of the considerable amount of numerical calculations involved.

Acknowledgment

We are indebted to D. Ridder for skillfully preparing the *p-i-n* structures.

Discontinuities in the Center Conductor of Symmetric Strip Transmission Line*

H. M. ALTSCHULER† AND A. A. OLINER†

Summary—A systematic measurements program has been carried out to check the validity of theoretical formulas for the equivalent circuit parameters of a variety of discontinuities in the center conductor of symmetric strip transmission line. These theoretical formulas have been in part previously available and are in part new or modified. Results indicate that, in general, these formulas are adequate for most engineering purposes and that certain of the network parameters can be neglected.

I. Introduction

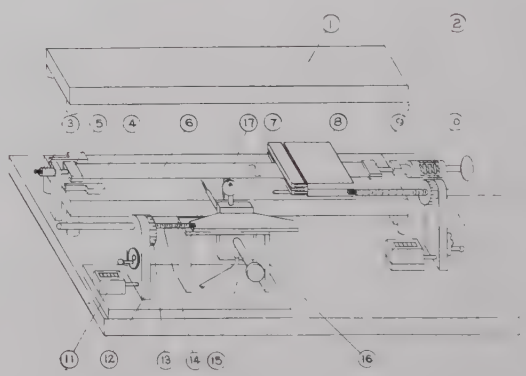
THEORETICAL expressions for the equivalent circuit parameters of a variety of discontinuity structures in the center conductor of strip transmission line have been available[1] for some time now. Many of these are simple, first order results derived from known formulas for the equivalent circuit parameters of related microwave structures. While some of these theoretical expressions had been compared with measurements in the past,[2,3] many had not been checked. The work described here, which was initiated some years ago, has been a systematic effort to fill this gap. It has thrown light on the accuracy of the formulas in question and, in addition, has resulted in the modification of certain of these formulas (Section III-I) and in the derivation of expressions not available heretofore (Sections III-B, D, and G).

The emphasis is placed on the comparison between theory and measurement. The many techniques and details involved in carrying out the construction of the center strips, in the measurement techniques, and in the network considerations required to cast the measured data into usable form are largely omitted. These points have been adequately discussed in a thesis.[4]

A brief description of the apparatus used, the nature of the center strip, and some comments on the precision employed follow immediately below. The balance of the report, after brief general remarks, is devoted to the comparison of theory with measurements for a variety of discontinuity structures. Each type of discontinuity is treated in a separate self-contained section which includes the theoretical formulas used, an indication of their derivation, comments on the representation employed, graphs comparing theory and measurement, and a discussion of the results obtained.

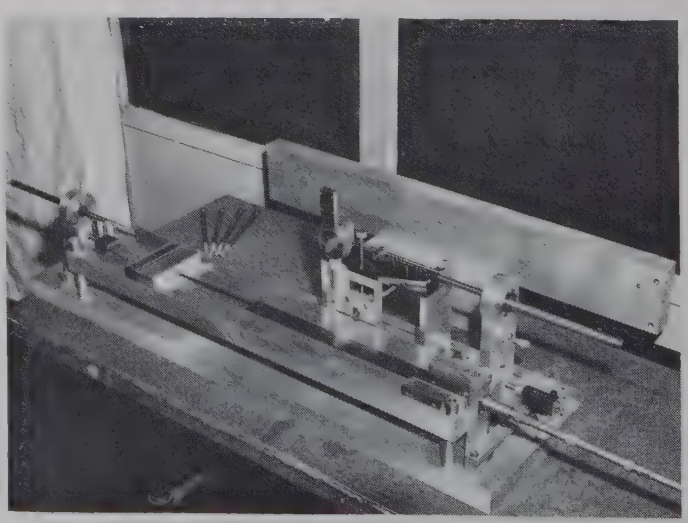
II. The Measurement Apparatus

In outline, the measurement equipment is a standard impedance measuring setup consisting of a source, a standing wave indicator followed by the discontinuity, which in turn is terminated in a variable short circuit. All components but the source were incorporated into a single unit [shown in Fig. 1(a) and (b)] which is



1	UPPER GROUND PLATE	5	INPUT CLAMP	9	END CLAMP	13	VERNIER
2	SUPPORT	6	STRIP	10	TENSION SPRING	14	DRIVING SCREW
3	INPUT CONNECTOR	7	DISCONTINUITY	11	REVOLUTION COUNTER	15	PROBE
4	MODE SUPPRESSOR	8	CHOKE SHORT CIRCUIT	12	GEAR TRAIN	16	STANDING WAVE METER
						17	DIAL GAUGE

(a)



(b)

Fig. 1—Measurement apparatus. (a) Sketch of front view. (b) Photograph of back view.

* Manuscript received by the PGMTT, November 24, 1959; revised manuscript received, January 18, 1960. The research reported was conducted under Contract No. AF-19(604)-2031 sponsored by the Air Force Cambridge Res. Center, Air Res. and Dev. Command.

† Microwave Res. Inst., Polytechnic Institute of Brooklyn, Brooklyn 1, N. Y.

[1] A. A. Oliner, "Equivalent circuits for discontinuities in balanced strip transmission line," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 134–143; March, 1955.

[2] A. D. Frost, private communication, Tufts University, Medford, Mass.; April 30, 1954.

[3] H. S. Keen, "Scientific Report on Study of Strip Transmission Lines," Airborne Instruments Lab., Mineola, N. Y., Rept. No. 2830-2; December, 1955.

[4] M. S. Stillman, "Measurement of Discontinuities in Symmetric Strip Transmission Line," M.E.E. thesis, Polytechnic Institute of Brooklyn, N. Y.; June, 1958.

described here to indicate the care with which the measurements were carried out.

The equipment is unique in that its center conductor is a flat, thin strip which is suspended in tension between an input connector and a spring-loaded clamp at the other end of the ground plates. This feature was chosen so that the center strip can be completely surrounded by air. These conditions correspond closely to the assumptions underlying the theoretical formulas that the center strip is of zero thickness and is immersed in a single homogeneous medium. An operating frequency of 1500 mc and a characteristic impedance of 50 ohms with a center conductor width of 1.5 inches and a plate spacing of 1.051 inches were chosen. Such large dimensions were employed to reduce mechanical problems and to increase accuracy.

The equipment is shown from two different views in Fig. 1(a) and (b). The upper ground plate is shown in an exploded view. All the equipment is mounted on a sturdy base plate. The two carefully machined aluminum ground plates are supported by cylindrical spacers. The center conductor of the input connector ends in an input clamp which has been made as small as possible to keep the associated discontinuity reasonably low. An interchangeable center strip, which includes the discontinuity, runs from the input clamp to the end clamp. The two mode suppressor plates near the input discriminate strongly against the radiating (parallel plate) TEM mode in that the dominant rectangular guide mode in this region is well beyond cutoff. The standing-wave meter incorporates a side probing arrangement similar to that used by Cohn.[5] It is driven by a gear train with an associated revolution counter and a wheel vernier with a ± 0.0001-inch readability. A conventional probe, modified by extending the probe wire and its shield by several inches, was used. The probe can also be moved in a vertical direction with respect to the probe carriage, and its distance from the bottom ground plate can be measured by means of a dial gauge. The standing-wave meter and the associated equipment are mounted on a small separate base plate which can be rotated slightly by means of adjustment screws in order to bring the probe travel into parallelism with the center strip. The movable choke-type short circuit has associated with it the same type of drive and counter mechanism as the standing-wave meter. It is built in the form of a "sandwich" about the center strip and is guided by a groove in the bottom ground plate. The top "half" of the short circuit is removable to allow the center conductor to be changed. The end clamp holds the center strip and transmits the force of the (variable) tension spring to the center strip.

Center strip construction was originally the subject of much experimentation with the object of reducing or eliminating undesirable deformations which occur in the neighborhood of the discontinuity when the strips are in tension. The construction finally arrived at includes a 0.001-inch "Mylar" polyester film (no discontinuity is cut in this) sandwiched between two 0.001-inch copper foil strips. After assembly with a tacky teflon cement, the total thickness is about 0.0035 inch. In order to minimize deformation, tension in the copper is relieved by certain cuts made near the discontinuity (perpendicular to the direction of force) across the copper strips. Before measurements were made, the strips were stress relieved in tension for a number of hours. Both "cutting" and photo-etching techniques were employed in constructing discontinuities.

The over-all performance of the measurement setup was satisfactory. Excessive radiation was guarded against by careful strip construction techniques and by centering the strip between the plates with utmost care. Inevitably, there remained a very small amount of radiation which did not appear to influence the measurements at all. The quality of the measurements is reflected by such typical insertion VSWR values as $|\gamma| = 1.167 \pm 0.006$, $|\gamma| = 1.422 \pm 0.008$ and $|\gamma| = 148 \pm 1.5$. The curve of Fig. 2 shows the plotted data ($D+S$ vs S) for a typical step discontinuity together with a plot of the scatter (ΔD vs S) which remained after the network parameters were abstracted by analytical means. To emphasize how small the scatter is in relation to the data, $D+S$ and ΔD are shown to the same scale. In the region of $|\gamma|$ near 1.5 the uncertainties ranged from 0.006 to 0.010; in the same region of $|\gamma|$ the uncertainties associated with precision X-band equipment used to measure discontinuities in rectangular waveguide ranged from 0.002 to 0.007. This is viewed as an indication of very good precision when the construction and mounting difficulties are kept in mind. The errors associated with the necessary reference plane measurements were unfortunately considerably larger. These are thought to have been of the order of ± 0.020 inch, *i.e.*, about $\pm \lambda/400$. These errors, though not serious, made themselves quite evident in parameters which were sensitively dependent on reference plane location.

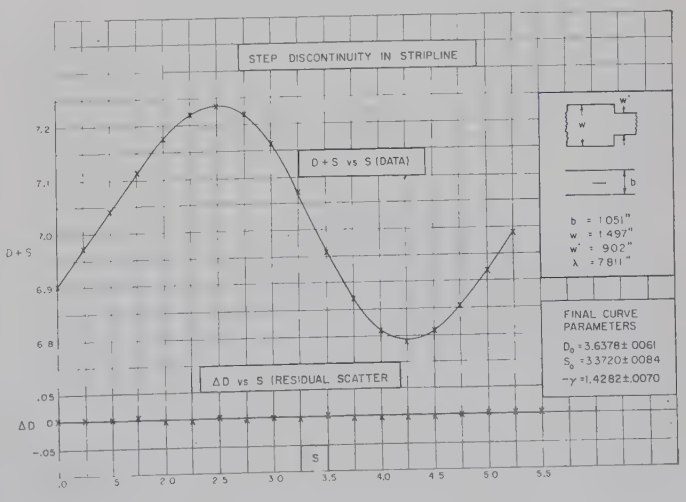


Fig. 2—Typical data curve and associated residual scatter.

[5] S. B. Cohn, "Problems in strip transmission line," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 119–126; March, 1955.

III. Comparison Between Theory and Measurement

A. General Remarks

The measurements described in the following were all carried out at very nearly 1500 mc, *i.e.*, at a wavelength of about 7.87 inches, except for the abruptly-ended center conductor (Section III-B) for which measurements were carried out at various frequencies. The cross section of the strip transmission line used is shown in Fig. 3, in which b, w, and t are implicitly defined. The plate spacing b was maintained accurately throughout the measurements, but the strip width w actually varied slightly from strip to strip. These variations were always properly taken into account. Strip thickness t was taken to be zero in all cases but in the calculations of the characteristic impedance ratios of step discontinuities (Section III-F), where this ratio and the measured insertion VSWR were very nearly equal to each other.

The characteristic impedance values used in certain formulas below were obtained from the following known expressions:

$$Z_o = \frac{30\pi\left(1 - \frac{t}{b}\right)}{D/b}\,; \quad Z_{02}/Z_{01} = D_1/D_2 \tag{1}$$

where, for small strip thickness t, D is given accurately by

$$D = bK(k)/K(k') + \frac{t}{\pi}\left[1 - \ln\left(\frac{2t}{b}\right)\right] \tag{2}$$

and, to a very good approximation, for $w/b > 0.5$, by

$$D = w + \frac{2b}{\pi}\ln 2 + \frac{t}{\pi}\left[1 - \ln\left(\frac{2t}{b}\right)\right]. \tag{3}$$

Here k and k' are defined by

$$k = \tanh(\pi w/2b), \qquad k' = (1 - k^2)^{1/2}. \tag{4}$$

In the preceding expressions, D is the "equivalent strip width," *i.e.*, the width of a strip with an associated *uniform* field (without fringing) but with the same capacitance as the actual strip.

Unless stated otherwise, all impedances and reactances are assumed to be properly normalized to characteristic impedance. In other words, the normalized impedance of a matched termination is unity. Normalization is indicated by a prime. All dimensions and lengths are given in inches.

Since the equivalent circuit parameters are expressed in normalized form, these circuits are to be used in the following usual fashion: Impedances which terminate the network must be taken as normalized; any input impedance computed from the network is then automatically normalized. Measured parameters are automatically expressed in normalized terms unless characteristic impedance values are explicitly introduced into the computations which transform the measured data into a network representation.

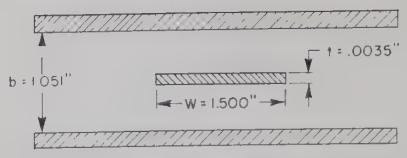


Fig. 3—Cross section of strip transmission line.

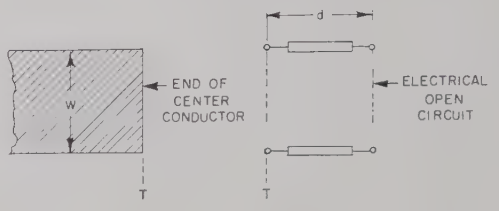


Fig. 4—Abruptly-ended center conductor. (a) Physical structure. (b) Equivalent circuit.

B. Abruptly-Ended Center Conductor

The physical structure and the equivalent network of the abruptly-ended center conductor are shown in Fig. 4. Since, to a crude approximation, one can expect an electrical open circuit to be located at the plane T, the representation chosen results in relatively small values of the transmission line length d for practical strip transmission lines.

The only rigorous theoretical result available for this structure is the known static value for the center conductor of infinite width (fringing at an edge). In this case one has a value of d equal to c where

$$c = \frac{b \ln 2}{\pi}\,. \tag{5}$$

For the case in which the center conductor width w is not infinite, a theoretical expression based on both edge and corner fringing has been developed. The edge contribution is based on the above (infinite width) expression, while the corner contribution has been obtained empirically from measured data since no theoretical solution for the capacitance of a corner between parallel plates is available. The formula is

$$d = \frac{1}{\kappa}\cot^{-1}\left[\frac{4c + 2w}{c + 2w}\cot(\kappa c)\right], \qquad \kappa = \frac{2\pi}{\lambda} \tag{6}$$

where λ is the wavelength in the medium. For most practical dimensions (κc small), one can approximate (6) by (7) which is seen to be independent of κ, *i.e.*, of both the frequency and the dielectric constant of the material contained in the transmission line. The simplified formula is

$$d = c\left(\frac{c + 2w}{4c + 2w}\right), \tag{7}$$

which approximates (6) to within 3 per cent for $\kappa c \leq 0.3$.

Eqs. (6) and (7) already hold implicit the value of the empirically obtained corner fringing capacitance C_{cf}. An independently derived formula for d/b, which explicitly contains the corner fringing capacitance as a parameter, has been available[6] for the case of parallel-coupled strips, one of which is open-ended. In the limit in which the strips become completely uncoupled this expression reduces exactly to the simplified form (7). The symbol C_f'' employed by these authors is defined as one-half of C_{cf}; their empirical value for $2C_f''$ in micromicrofarads for a zero-thickness center strip is $0.019\,\epsilon_r b$, where b is in inches and ϵ_r is the relative dielectric constant. The corresponding value of C_{cf} implicit in (7) is $0.011\,\epsilon_r b$.

Theoretical and experimental values for d, the location of the electrical open circuit with respect to the end of the center conductor, are compared in Fig. 5, in which both strip width (the parameter) and d have been normalized to the plate spacing b. Measurements were made at six different wavelengths ranging from $\lambda = 5.4$ inches to $\lambda = 11.8$ inches. The "bars" crossing the theoretical curves result from both experimental scatter and the small variation of d/b with frequency. The theoretical curve is based on (7). The writers have also been informed[7] that data taken at the Stanford Research Institute and the Airborne Instruments Laboratory over a portion of the ω/b range lie near the upper limit of the vertical scatter lines shown in Fig. 5.

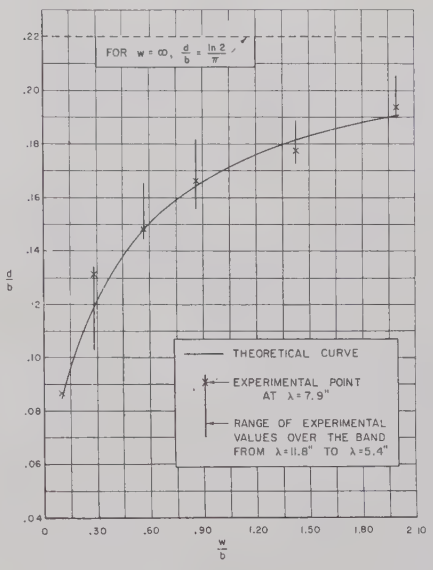


Fig. 5—Location of equivalent open circuit for abruptly-ended center conductor.

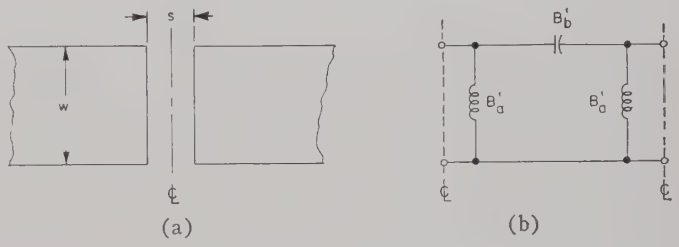


Fig. 6—Gap in center conductor. (a) Physical structure. (b) Equivalent circuit at centerline reference planes.

C. Gap in Center Conductor

The physical gap structure and its equivalent admittance pi network[8] at centerline reference planes are shown in Fig. 6.

The gap formulas[1] were originally derived from available results for E-plane slit-coupled rectangular waveguides. The derivation proceeded by first obtaining an approximate parallel plate waveguide model of the actual discontinuity structure, consisting of two parallel plate waveguides coupled by an infinite slit (gap) of width s, and then by simply employing the known rectangular waveguide result by re-expressing it in the limit as the guide width becomes infinite. As a result of the manner of derivation, the strip width w does not enter as a parameter into the formulas obtained. As w/b becomes very small the formulas can consequently be expected to deteriorate. The expressions are

$$B_a' = -\frac{2b}{\lambda}\ln\left[\cosh\left(\frac{\pi s}{2b}\right)\right], \tag{8}$$

$$B_b' = \frac{b}{\lambda}\ln\left[\coth\left(\frac{\pi s}{2b}\right)\right]. \tag{9}$$

Comparison of theoretical results with measured data furnished by the Airborne Instruments Laboratory was made previously[1] for the case of a 50-ohm line and showed rather good agreement over a wide range of frequencies. Subsequent measurements[9] indicate that for very narrow strips (high characteristic impedances) the values of B_a' and B_b' decrease noticeably. In extreme cases the values predicted by these simple formulas may be too large by a factor of 1.4 for B_a' and 2.0 for B_b'.

D. Rectangular Slot in Center Conductor

The physical slot structure and its equivalent admittance pi network are shown in Fig. 7. The slot is narrow and is centered on the strip conductor; when $d = w$, *i.e.*, when the slot runs completely across the strip, the structure is called a gap. It is known that a narrow gap is capacitive while a short slot is inductive, so that the slot should be resonant for some intermediate length.

[6] S. B. Cohn, *et al.*, "Research on Design Criteria for Microwave Filters," Stanford Res. Inst., Menlo Park, Calif., pp. 115–116; June, 1957.

[7] S. B. Cohn, private communication.

[8] The reviewers point out that many of the inductive network elements, such as those in Figs. 6 and 7, are actually negative capacitive elements and that therefore it is proper to use capacitor rather than inductor symbols in these instances. The authors agree that, to a first order, the formulas for these elements have the frequency dependence of negative capacitance. Nevertheless, the customary symbolism has been retained in this paper for the following reasons: 1) It is not clear whether such network symbols should imply primarily frequency dependence or sign; 2) at any given single frequency the sign is of course more important; 3) if an element is to be used over a frequency band, the detailed expression for the element must be carefully examined in any case, since not all elements possess a simple frequency dependence.

[9] Keen, *op. cit.*, Figs. 18 and 19.

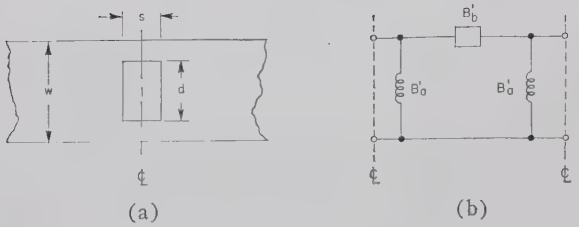


Fig. 7—Slot in center conductor. (a) Physical structure. (b) Equivalent circuit at centerline reference planes.

These remarks apply to the series arm B_b' of the pi equivalent network for the slot. The inductive shunt arms are due to the finite width of the slot, but their effect is small if the slot is narrow.

The hitherto unpublished formulas given below for a slot in the center conductor were derived by Suzuki.[10] Their evaluation is based on the use of the following simple model for the strip transmission line. By replacing the fringing field at the strip edges by an extension of the strip width which is terminated by an open circuit, the strip transmission line is transformed into two parallel plate guides of finite width placed back to back and coupled by the slot. By duality considerations, the geometry is then related to a flat metal rectangle located parallel to the electric field in a parallel plate waveguide of width equal to the *height* of the original strip line. The equivalent circuit for the metal rectangle in parallel plate guide is then obtained approximately from an accurately derived result for a tuned post[7] in rectangular waveguide.

The formulas for the slot are given below. In the limit as the slot runs completely across the center conductor, these complicated formulas reduce exactly to the expressions for the gap given in Section III-C.

$$B_a' = \frac{-\frac{2b}{\lambda}\ln\cosh\frac{w\tau}{b}}{1+P} \tag{10}$$

$$B_b' = -\frac{b}{\lambda}\ln\sinh\frac{w\tau}{b} - \frac{B_a'}{2} - \frac{b\lambda}{w^2}\left(1+\frac{2b}{\pi w}\ln 2\right)Q \tag{11}$$

where

$$Q \simeq \frac{1}{4\tau}\left[\frac{-1}{\ln\beta} - \left(\frac{1-\beta^2}{\ln\beta}\right)^2 \sum_{n=1}^{N} f(2n\tau)\frac{X_n^2}{n}\right].$$

N is the integer nearest the quantity $(0.7/\tau - 1)$,

$$P \simeq \frac{1}{4}\left(\frac{1-\beta^2}{\ln\beta}\right)^2 \sum_{n=1}^{5} g(2n\tau)\frac{X_n^2}{n^2} + \tau^2 Q,$$

and where

$$\tau = \frac{\pi s}{2w} \quad \text{and} \quad \beta = \cos\frac{\pi d}{2w}.$$

The quantity X_n is given by

$$X_1 = 1,\ X_2 = -1 + 3\beta^2,\ X_3 = 1 - 8\beta^2 + 10\beta^4,$$
$$X_4 = -1 + 15\beta^2 - 45\beta^4 + 35\beta^6,$$
$$X_5 = 1 - 24\beta^2 + 126\beta^4 - 224\beta^6 + 126\beta^8.$$

The functions $f(x)$ and $g(x)$ are plotted in Fig. 8. The formulas for B_a' and B_b' are applicable only when $\tau \geq 0.15$ and $d/w \geq 0.25$.

Theoretical and experimental values for the parameters of slots and certain gaps are compared in Fig. 9. It can be seen that these values are in good agreement with each other. The series susceptance B_b' is the dominant parameter in the representation for the slot since the associated values of B_a' are quite small. It is of interest to note that the series susceptance B_b' for the slots considered are all inductive, while B_b' for gaps are capacitive. The curve for $s = 0.155$ shows how rapidly the value of B_b' passes through resonance and that resonance occurs when the slot is "almost a gap." This is well confirmed by the experimental data. It is consequently recognized that a resonant slot is an impractical structure. Quite apart from the high sensitivity of the resonant point to the exact slot dimensions, the slot side walls $(w - d/2)$ would have to be exceptionally thin to obtain resonance and would consequently present virtually insurmountable construction problems.

E. Round Hole in Center Conductor

The physical structure and the equivalent pi network for the round hole in a strip line center conductor are shown in Fig. 10. It should be noted that this representation is at the centerline of this symmetric structure.

The theoretical formulas for B_a' and B_b' have been available for some time[1] and were obtained by means of variational expressions. The stored power terms (numerators) involved were obtained by small aperture considerations from known results for the E-plane aperture coupling of two rectangular waveguides. The formulas are

$$B_b' = -3\lambda b D/4\pi d^3, \tag{12}$$
$$B_a' = 1/4B_b'. \tag{13}$$

Extensive measurements on this discontinuity were carried out at the Airborne Instruments Laboratory, the results of which are already available.[1,3] The theoretical values were found to be in good agreement with the results of the measurements referred to.

F. Step (Change in Width) in Center Conductor

The physical structure and an equivalent network for the step (abrupt change of center conductor width) are shown in Fig. 11. Since both characteristic impedances

[10] M. Suzuki, "Circuit Parameters of a Tuning Post in a Rectangular Waveguide and its Applications," Microwave Res. Inst., Polytechnic Inst. of Brooklyn, N. Y., Rept. No. R-591-57, PIB-519; July, 1957.

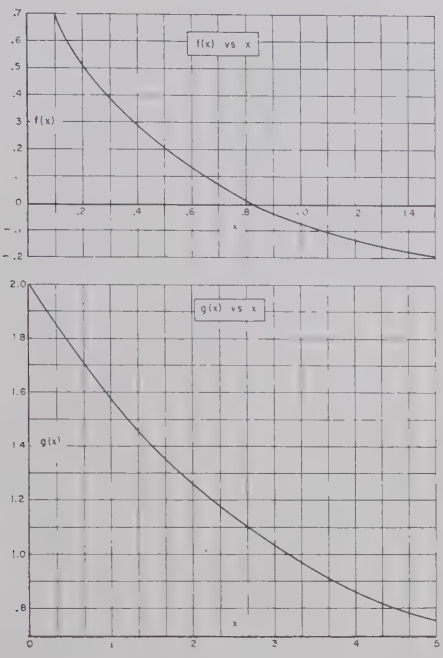


Fig. 8—Functions employed in slot formulas

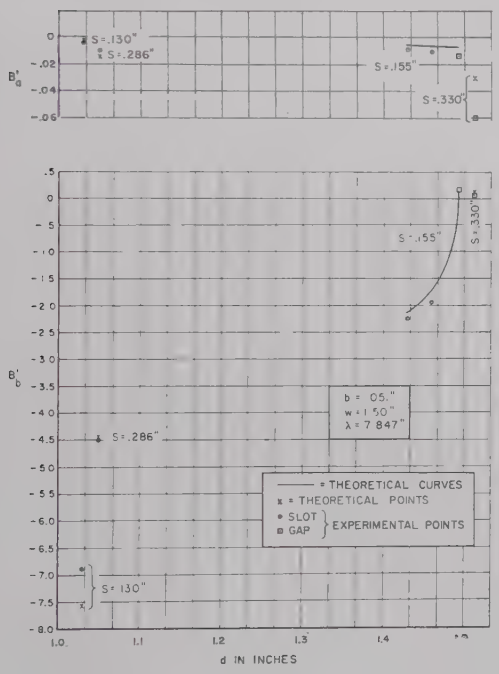


Fig. 9—Circuit parameters of slots and gaps.

have been normalized here by taking them as unity (see Section III-A), the ideal transformer $\sqrt{Z_{02}'}$:1 appears explicitly in the circuit representation. If the (unnormalized) characteristic impedances of the two transmission lines involved are defined as Z_{01} and Z_{02}, respectively, then

$$Z_{02}' = Z_{02}/Z_{01} \quad \text{and} \quad X' = X/Z_{01}.$$

The theoretical formulas (for X', l_1, and l_2) were obtained[1] via a Babinet equivalence procedure from the known parameters of a step in the height of rectangular

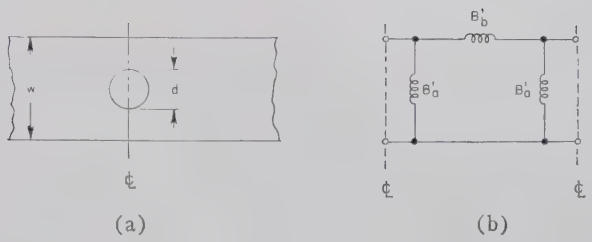


Fig. 10—Round hole in center conductor. (a) Physical structure. (b) Equivalent circuit at centerline reference planes.

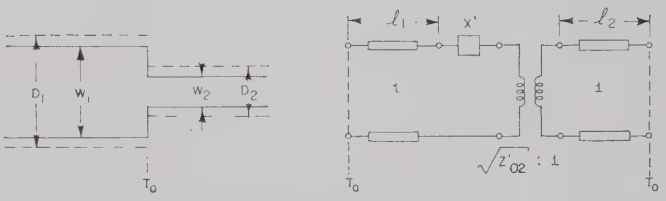


Fig. 11—Step (change in width) in center conductor. (a) Physical structure. (b) Equivalent circuit.

waveguide. The approximation used was such that l_1 and l_2 depend only on the strip-line ground-plate spacing. This simple dependence is certainly open to question, but no more exact formulation has been attempted. The theoretical formulas are

$$X' = \frac{2D_1}{\lambda} \ln \csc\left(\frac{\pi D_2}{2D_1}\right), \tag{14}$$

$$l_1 = -l_2 = \frac{b \ln 2}{\pi}. \tag{15}$$

The values for D_1, D_2, and Z_{02}' are obtained from (1) to (4). The theoretical insertion VSWR for the structure, r, is computed from X' and Z_{02} by the following formula:

$$r = a + \sqrt{a^2 - 1}, \quad \text{where } a \equiv \frac{X'^2 + Z_{02}'^2 + 1}{2Z_{02}'}. \tag{16}$$

Theoretical and experimental values for some of the circuit parameters are presented in Table I. The values

TABLE I

w_2 (inches)	Theoretical			Experimental r	Semi-Experimental (inches)	
	Z_{02}'	r	X'		l_1	l_2
.152	3.375	3.429	.408	3.348 ± .035	+.102	−.102
.310	2.577	2.614	.287	2.614 ± .025	+.063	−.067
.600	1.842	1.857	.142	1.825 ± .013	+.065	−.079
.902	1.436	1.440	.059	1.428 ± .007	+.042	−.065
1.205	1.174	1.175	.014	1.204 ± .018	−.047	+.056

given assume various wavelengths λ very close to 7.80 inches, strip widths w_1 nearly equal to 1.50 inches, and a ground plate spacing maintained at $b > 1.051$ inches. It follows that the theoretical values for the transmission

line lengths l_1 and l_2 are

$$l_1 = -l_2 = 0.232 \text{ inch.}$$

Above all, it will be noted that the theoretical values for Z_{02}' and r, and the experimental values of r, all fall very close to each other, differences between corresponding values being roughly of the order of 1 per cent. One concludes at once that the change in characteristic impedance (Z_{02}') dominates by far over the series reactance X' in determining the values of r. It follows as a corollary that, since experimental values of X' are obtained by the use of the expression

$$X' = \left[\frac{Z_{02}'}{r}(r^2+1) - (Z_{02}'^2+1)\right]^{1/2}, \qquad (17)$$

X' is exceptionally sensitive to the smallest errors in both Z_{02}' and r. This was indeed found to be the case, in fact so much so that the experimental values of X' were completely unreliable. They have consequently not been included here. The theoretical values of X' are tabulated, however. While the experiments performed have not verified these theoretical values, they do indicate that the orders of magnitude of X' are correct.

The experimental values of l_1 and l_2 must be derived from measured data by formulas which, among other quantities, involve the (previously determined) experimental value of X'. In view of the unreliability of the experimental values of X', the derived values obtained for l_1 and l_2 were equally unreliable. Computations made using the theoretical X' in conjunction with other data obtained by measurement gave values for l_1 and l_2 ("semi-experimental") which appear quite reasonable. From these values (which are tabulated) one concludes that l_1 is indeed approximately equal to l_2 in magnitude and of opposite sign as the theoretical model predicts. It is seen, however, that the magnitudes involved are considerably smaller than $l_1 \simeq 0.23$ inch, as predicted for this case by (15). In all but one case ($w_1 = 1.204$, the most difficult to measure in this regard) l_1 and l_2 are, respectively, properly negative and positive. In view of these results, it is recommended that (15) be disregarded, and that for most purposes l_1 and l_2 be taken as zero.

The conclusions to be drawn from the preceding comparison of theory and experiment are: The step in strip line is very well approximated at T_0 by only the transformer ($\sqrt{Z_{02}'}$:1); the series reactance X' is negligible for almost all practical purposes, and the transmission lines ($l_1 = -l_2$) are small and unimportant.

G. Junction of Strip Transmission Line and Multistrip Transmission Line

The term multistrip transmission line has been employed previously[11] to refer to an infinite number of parallel, equispaced strips of identical widths symmetrically located between two ground planes. In the present context the term is used in the same fashion except that a *finite* number (m) of strips, with $m>1$, will be understood. The structure of interest is the junction between such a multistrip line and a (single center conductor) strip line as shown in Fig. 12.

A multistrip line can support $m+1$ different TEM modes, all of which, of course, have the same guide wavelength, *i.e.*, the free-space wavelength in the medium. In a two-strip line, for example, ($m=2$), three modes can propagate with electric fields as shown in Fig. 13.

The only mode under consideration here is the mode for which the fields surrounding each strip are almost identical to the field of the dominant (nonradiating) mode in (single center conductor) strip line. Such a mode is illustrated in Fig. 13(b) and Fig. 14 for $m=2$ and $m=4$, respectively. To distinguish this mode from all other modes in multistrip line, it will be arbitrarily referred to here as the "A mode." The junction in question then is that between usual strip line and the "A mode" in multistrip line.

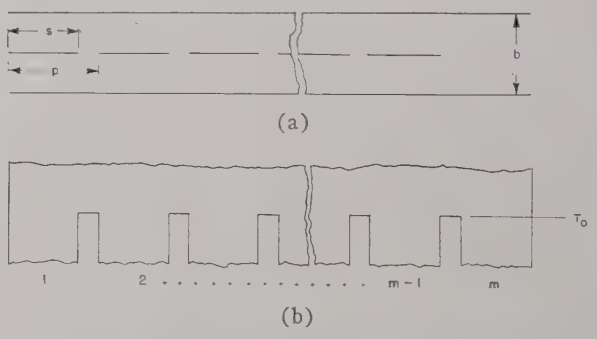


Fig. 12—Multistrip transmission line. (a) Cross section view. (b) Top view of center conductor.

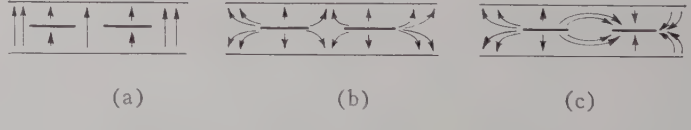


Fig. 13—Modes in a two-strip line. (a) Parallel plate (radiating) mode. (b) Even two-strip mode. (c) Odd two-strip mode.

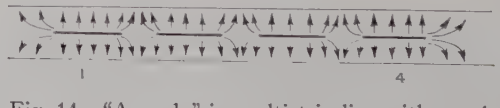


Fig. 14—"A mode" in multistrip line with $m=4$.

It can be justifiably argued that the equivalent circuit representation of the strip-to-multistrip A-mode junction consists primarily of a (change of characteristic impedance) transformer. The representation should also include a series reactance, a small reference plane shift from the physical junction plane, and some small loss parameter(s) due to mode conversion. Experience

[11] A. A. Oliner and W. Rotman, "Periodic structures in trough waveguide," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-7, pp. 134–140; January, 1959.

with the step in strip transmission line, however, strongly indicates that these additional parameters are quite small, so that the transformer alone at the plane T_o is an adequate representation for most purposes. This representation is shown in Fig. 15.

A previous result[11] for the case $m=\infty$ gives an expression for $Z_{o\infty}/Z_o$ based on a Babinet equivalence argument which uses the known susceptance of a capacitive slit in rectangular waveguide as a basis. This formula is

$$\frac{Z_{o\infty}}{Z_o} = 1 + \frac{2p}{\pi b} \ln \csc \frac{\pi s}{2p} . \tag{18}$$

By making use of this expression and the notion that the field in the multistrip line (m finite) is identical with the fields in the m infinite case everywhere except at the "outer" halves of strips 1 and m, one can derive a similar formula for m finite. The field associated with the two "outer" half-strips in this case can be regarded as the field of a (single) strip transmission line of width s. The resulting expression is

$$\frac{Z_{om}}{Z_o} = \frac{Z_{o\infty}/Z_o + A}{1 + A} , \tag{19}$$

where

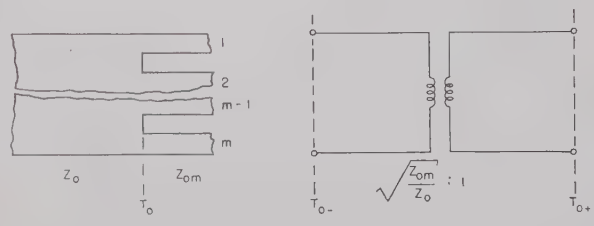
$$A = \frac{K(k')}{K(k)} \frac{p}{mb} ; \quad k = \tanh\left(\frac{\pi s}{2b}\right), \quad k' = \sqrt{1 - k^2}.$$

The function K is the complete elliptic integral of the first kind. It can be seen that in the limiting cases of $m=0$ and $m=\infty$, the ratio Z_{om}/Z_o properly reduces to unity and to $Z_{o\infty}/Z_o$.

The measurement of Z_{om}/Z_o can be accomplished with relatively good accuracy; in addition, A is a constant which depends only on the known geometry of the multistrip line. In view of the simple formula relating these quantities to $Z_{o\infty}/Z_o$, the latter can be derived readily from the experimental value of Z_{om}/Z_o. The quantity $Z_{o\infty}/Z_o$ is of particular interest in context with certain periodic structures.[8]

Measurements on a six-strip transmission line were carried out to determine the parameters of the function at T_o with the objective of deriving $Z_{o\infty}/Z_o$ from the measured Z_{om}/Z_o value. Six was thought to be a sufficiently large number to satisfy reasonably well the inherent assumption that all but strips 1 and m are fed by a uniform field. The strip was measured at a wavelength $\lambda=7.830$ inches; the various dimensions were $b=1.051$ inches, $s=0.136$ inch, and $p=0.272$ inch. An average of three precision measurements gave the results which are compared in Table II with the corresponding theoretical values. The results of the experiment also indicated, as one would expect, that both reference planes of the transformer are in fact located slightly to the left of T_o (see Fig. 15); the left reference plane by approximately 0.05 inch, the right by about 0.04 inch.

The formula for Z_{om}/Z_o for $m=2$ involves the characteristic impedance of the even two-strip mode shown in Fig. 13(b), for which an exact result is available in the literature.[12] In terms of the notation employed in the reference cited, $Z_{02}/Z_o = Z_{oe}/2Z_o$. Because the multistrip formula for Z_{om}/Z_o can be expected to be least accurate in the two-strip case, it was compared to the exact two-strip formula. This comparison is presented graphically in Fig. 16. Since even in this worst case the correspondence is relatively good, it is felt that the multistrip formula can be usefully employed when $m=3$ or 4 and that it should be quite accurate for $m>4$.



(a) (b)

Fig. 15—Junction between strip and multistrip transmission lines. (a) Top view of center conductor. (b) Simplified equivalent circuit.

TABLE II

	Z_{om}/Z_o	$Z_{o\infty}/Z_o$
Experimental	1.0535 ± 0.005	1.0588 ± 0.0055
Theoretical	1.0522	1.0573

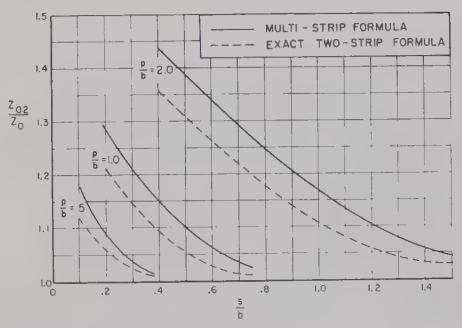


Fig. 16—Comparison between exact two-strip formula and multistrip formula for $m=2$ (worst case).

When a multistrip line is abruptly ended, an electrical open circuit will be located a small distance beyond the physical end as in the single center conductor case (see Fig. 4). Measurements were carried out on a multistrip line $m=6$, $b=1.051$ inches, $s=0.136$ inch, $p=0.272$ inch, at $\lambda=7.830$ inches to determine d, the distance from the end of the conductor to the electrical open circuit. The distance d was found to be 0.165 inch $\pm$ 0.015 inch. No theoretical expression for d is available. The number, however, falls roughly into the range one would expect by comparing it to the abruptly ended single strip case (see Fig. 5).

[12] S. B. Cohn, "Shielded coupled-strip transmission line," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 29–38; October, 1955.

H. Sharp Bends: Arbitrary Angle, Right Angle

The parameters of sharp bends in the center conductor of strip line were measured as a function of angle θ for a single strip width and again as a function of strip width w for a single angle (90°). These two sets of measurements are presented essentially separately below.

The physical structure of the *arbitrary angle bend* and two alternative equivalent circuits are shown in Fig. 17.

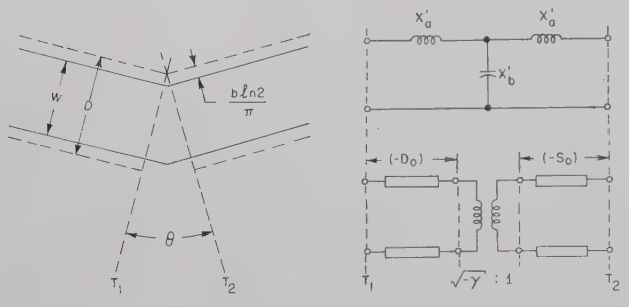


Fig. 17—Sharp bend of arbitrary angle in center conductor. (a) Physical structure. (b) Two alternative equivalent circuits.

The reactance tee network involves the normalized reactances X_a' and X_b' which are obtained directly from theory. The theoretical parameters of the "tangent" (or transformer) network, however, are based on the values of X_a' and X_b'. The tangent network is a more useful representation for a bend since $-\gamma$ equals the insertion VSWR of the structure. It should be noted that reference planes T_1 and T_2, which are perpendicular to the strip on the two sides of the bend, intersect at the inside corner of the equivalent width strip (width D) rather than at the inside corner of the physical strip itself.

The theoretical formulas for X_a' and X_b' were derived[1] from the results for an E-plane bend in parallel plate waveguide by means of a Babinet equivalence procedure. These formulas are

$$X_b' = -\frac{\lambda}{2\pi D}\cot(\theta/2), \tag{20}$$

$$X_a' = \frac{2D}{\lambda}\left[\psi(x) + 1.9635 - \frac{1}{x}\right], \tag{21}$$

where, with θ in degrees,

$$x = \frac{1}{2}\left(1 + \frac{\theta}{180}\right), \qquad \frac{1}{2} < x < 1.$$

The function $\psi(x)$ is tabulated.[13]

[13] E. Jahnke and F. Emde, "Table of Functions," Dover Publications, Inc., New York, N. Y., p. 16; 1945.

The expressions required for abstracting the parameters of the tangent network (γ, D_o, and S_o) from those of the tee are

$$-\gamma = C + \sqrt{C^2 - 1}, \tag{22}$$

where

$$C = \frac{1 + 2(X_a' + X_b')^2 + X_a'^2(X_a' + 2X_b')^2}{2X_b'^2},$$

$$\kappa D_o = \tan^{-1}\alpha, \tag{23}$$

where

$$\alpha = \frac{\gamma - X_a'(X_a' + 2X_b')}{(X_a' + X_b')(1 + \gamma)},$$

$$\kappa S_o = \kappa D_o \pm \pi/2. \tag{24}$$

For θ small ($c \simeq 1$),

$$-\gamma \simeq 1 - \left(2X_a' + \frac{1}{X_b'}\right), \qquad \alpha \simeq \frac{1}{X_b'} - 1. \tag{25}$$

Theoretical and experimental results are compared in Figs. 18 and 19. The solid curves represent the theoretical results as computed from the formulas already given, *i.e.*, based upon the theoretical expressions for the parameters of E-plane bends in rectangular waveguide. The dashed curve (for $\theta > 60°$) for X_a' on the other hand is based upon *experimentally* obtained values for the parameters of E-plane bends in rectangular waveguide. Part of the discrepancy between (solid line) theory and experiment is consequently due to errors existing already in the theoretical result (X_a' only) for the rectangular waveguide bend. The discrepancy between (strip line) experiment and the dashed curve is somewhat larger. This is ascribed to an imperfection in the Babinet equivalent model for the bend in the following respect. In contrast to the Babinet model, which has impenetrable walls, the actual strip line possesses a fringing field which permits an additional small interaction to occur between the center conductor on one side of the bend and that on the other. This effect, which the Babinet model cannot account for, becomes more marked as the bend angle θ becomes larger.

For the small bend angles ($\theta < 30°$) the expressions for abstracting γ and D_o from X_a' and X_b' become quite sensitive and can lead to unreasonably large computational errors. The small angle approximations of these formulas, however, are very good, and in fact are substantially more reliable than the more "exact" expressions in this range. The theoretical values for γ and D_o obtained by their use are shown in the form of dot-dash curves. The $-\gamma$ vs θ curve is especially interesting in that it shows very graphically that it is advantageous to employ two or three smaller bends in place of one larger angle bend. For example, consider the experimental points at $\theta = 30°$, 45°, and 90°. The 90° bend has an insertion VSWR of 1.75; two 45° bends in tandem

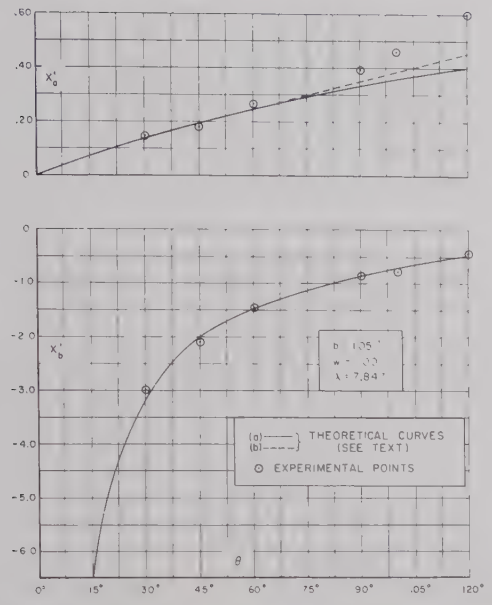


Fig. 18—Reactance tee network parameters for sharp angle bends.

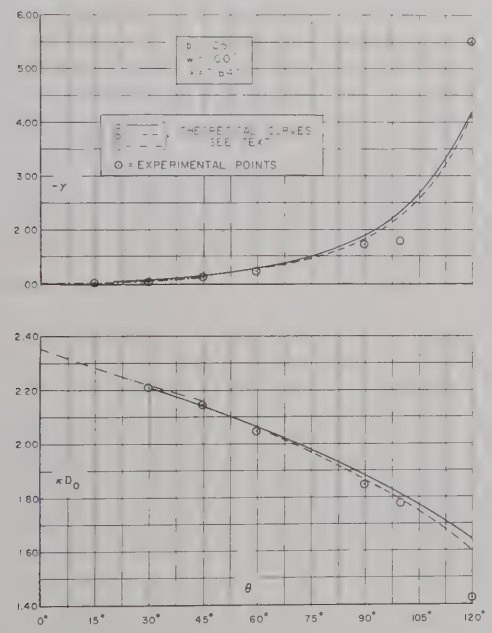


Fig. 19—Tangent network parameters for sharp angle bends

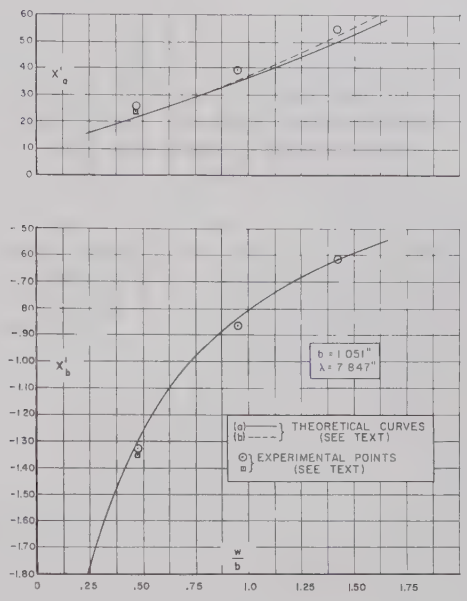


Fig. 20—Reactance tee network parameters for right-angle bends.

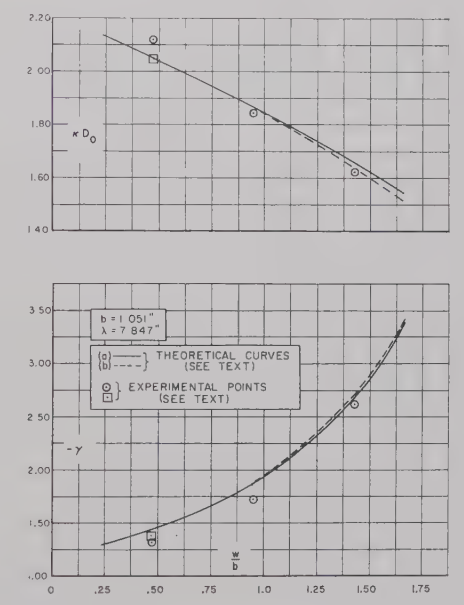


Fig. 21—Tangent network parameters for right-angle bends.

(and "far" from each other) can have a maximum insertion VSWR of 1.25; three 30° bends can have a maximum insertion VSWR of 1.15.

The *right-angle bend* is, of course, a special case of the arbitrary angle bend insofar as the physical structure and its equivalent circuit are concerned, *i.e.*, $\theta = 90°$ (see Fig. 17). The theoretical formulas have again been obtained[1] via a Babinet equivalence procedure, in this case from those for a right-angle E-plane bend in rectangular waveguide. The right-angle rectangular waveguide bend formula used as a basis is *not* a special case of the (rectangular waveguide) arbitrary angle bend formula, but is a more exact solution obtained especially for the 90° case. The right-angle bend formulas for X_a' and X_b' are given below. From these, the parameters γ, D_o, and S_o are, of course, abstracted as before by the use of (22) to (25).

$$X_a' = \frac{D}{\lambda}\left[1.756 + 4\left(\frac{D}{\lambda}\right)^2\right], \tag{26}$$

$$X_b' = 0.0725\left(\frac{D}{\lambda}\right) - 0.159/(D/\lambda). \tag{27}$$

Curves comparing theory and experiment are given in Figs. 20 and 21. As before, the solid line is based on

parameters for rectangular waveguide bends obtained by theoretical means and the dashed curves are the corresponding parameters obtained via rectangular waveguide measurements. As one would expect, the latter are in slightly better agreement with strip transmission line measurements than the strictly theoretical curves. In consequence of certain limitations of the measuring apparatus it was not possible to measure a single bend directly. Instead, either three or four (not necessarily similar) bends in tandem were measured at a time and the parameters for a single unknown bend were abstracted from the over-all data. The two experimental points shown for $w/b=0.475$ were obtained from two structures each consisting of four similar bends; however, the structures differed in that they involved different spacings between individual bends. The agreement between these two experimental points can be considered to be good in view of the many complications involved in the construction, measurement, and analysis of such four-bend structures.

I. Symmetric Tee Junction

The physical structure and the equivalent circuit for the symmetric tee junction are shown in Fig. 22. The symmetric arms of the tee structure are represented at the center line and the reactances X_a' and X_b' are normalized to Z_{01}. The representation of a symmetric lossless three-port in general requires four independent network parameters. In the representation employed here, one of these parameters is the length l of the transmission line which connects reference planes T_3 and T_3'. The approximate theory employed predicts, however, that l has a zero value. In consequence, the measured value of l directly expresses the error, if any, in the choice of T_3 as the plane at which the representation includes only the parameters X_a', X_b', and n.

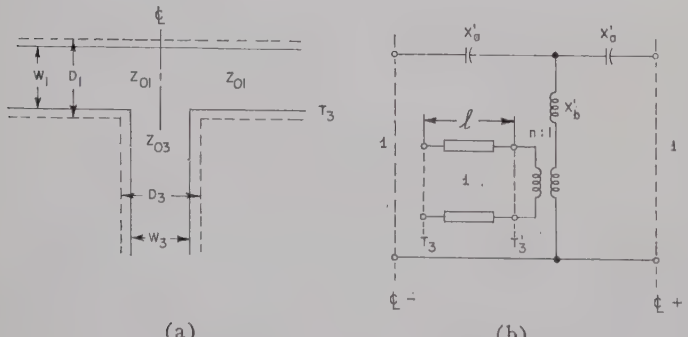


Fig. 22—Symmetric tee junction. (a) Physical structure. (b) Equivalent circuit.

The theoretical formulas (for X_a', X_b', and n) which have been available[1] were found to be adequate for X_a' and n. New expressions for X_b' have been derived, however, since the earlier formula had a very limited range of application. The present (as well as the earlier) formulas have been obtained via a Babinet equivalence procedure using known results for the parameters of the E-plane rectangular waveguide tee as a basis. They are

$$n' = \frac{\sin(\pi D_3/\lambda)}{(\pi D_3/\lambda)}; \qquad n = n'\sqrt{\frac{D_3}{D_1}}, \tag{28}$$

$$X_a' = -\frac{D_3}{\lambda}[0.785n]^2, \tag{29}$$

$$l = 0, \tag{30}$$

$$X_b' = -\frac{X_a'}{2} + \frac{1}{(n')^2}\left\{\frac{B_t}{2Y_o} + \left(\frac{2D_1}{\lambda}\right)\cdot\left[\ln 2 + \frac{\pi D_3}{6D_1} + \frac{3}{2}\left(\frac{D_1}{\lambda}\right)^2\right]\right\} \quad \text{for } \frac{D_3}{D_1} < 0.5, \tag{31}$$

where

$$\frac{B_t}{2Y_o} = \left(\frac{2D_1}{\lambda}\right)\left[\ln \csc\left(\frac{\pi D_3}{2D_1}\right) + \frac{1}{2}\left(\frac{D_1}{\lambda}\right)^2 \cos^4\left(\frac{\pi D_3}{2D_1}\right)\right],$$

$$X_b' = -\frac{X_a'}{2} + \frac{2D_1}{(n')^2\lambda}\left[\ln\left(\frac{1.43D_1}{D_3}\right) + 2\left(\frac{D_1}{\lambda}\right)^2\right] \quad \text{for } \frac{D_3}{D_1} > 0.5. \tag{32}$$

Seven tee structures with identical main line center strip widths, but with different stub line center strip widths, were measured. Measurements on strip line tee junctions have also been made at the Stanford Research Institute[14] with the objective of obtaining a compensated tee structure by empirical means. Their results have not been compared with the above formulas since a different equivalent circuit was employed by them.

Comparison of theory and measurement is shown in Figs. 23 and 24. The various parameters are plotted against the ratio of "main line" characteristic impedance to "stub line" characteristic impedance (Z_{01}/Z_{03}). This ratio equals D_3/D_1, where D_1 and D_3 are, respectively, the "equivalent widths" of the main line and stub line center conductors [see (1) to (4)]. As can be seen, the agreement for n is excellent. The series reactance X_a' is quite small and consequently relatively unimportant; the rather large percentage discrepancy between theory and experiment for X_a' can be expected in view of the difficulty of measuring such small parameters in a com-

[14] S. B. Cohn, *et al.*, "Design Criteria for Microwave Filters and Coupling Structures," Stanford Res. Inst., Menlo Park, Calif., Tech. Rept. No. 3; April 1–July 1, 1958.

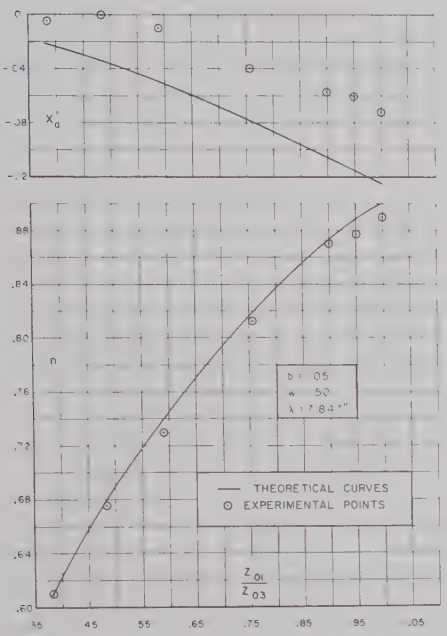


Fig. 23—Two of the circuit parameters for the symmetric tee junction.

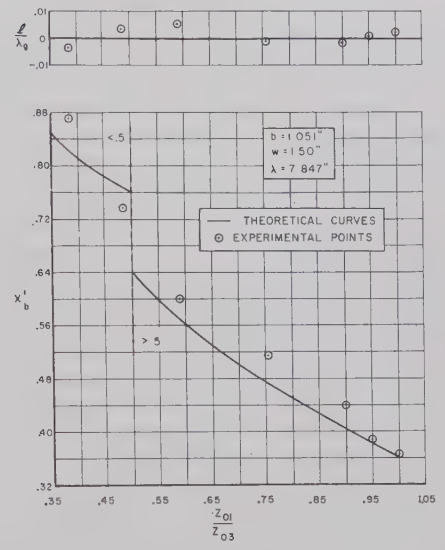


Fig. 24—The remaining two circuit parameters for the symmetric tee junction.

plicated structure in the presence of more significant parameters. The agreement for X_b' is actually much better (5 per cent or less) than a first glance at the curve would indicate. The two sections of the X_b' curve were computed by the formulas appropriate to the respective ranges. The experimentally determined length of line l (which is presented in Fig. 24 as normalized to guide wavelength) is seen to be sufficiently small to confirm the corresponding theory, *i.e.*, the terminals of the transformer are located so close to T_3 that line l is not required in the representation.

Acknowledgment

The authors wish to acknowledge the participation of M. Stillman, H. Hanft, S. Chertoff, and L. Horowitz in various phases of this work.

A Variational Integral for Propagation Constant of Lossy Transmission Lines*

ROBERT E. COLLIN†

Summary—**By assuming that the current on a lossy transmission line flows in the axial direction, only a variational integral for the propagation constant can be readily obtained. This variational integral shows that the usual power loss method of evaluating the attenuation constant is valid for general transmission lines. This variational integral also shows that the perturbation of the loss-free phase constant is due to the increase in magnetic field energy caused by penetration of the field into the conductors.**

* Manuscript received by the PGMTT, November 25, 1959; revised manuscript received, January 20, 1960. This work was supported in part by Air Force Cambridge Res. Center Contract AF19(604)3887.

† Elec. Engrg. Dept., Case Institute of Technology, Cleveland, Ohio.

Introduction

THE dominant mode of propagation on a loss-free transmission line is a TEM wave. In the transverse plane both the electric field and magnetic field may be derived from the gradients of suitable scalar functions of the transverse coordinates. The current flows entirely in the axial direction. Practical lines have finite conductivity and hence finite losses. As a consequence, there must be a component of the Poynting vector directed into the conductors and this in turn implies at least a longitudinal component of electric field. In general, longitudinal components of

both electric and magnetic fields will exist (some exceptions are the coaxial line, single wire line, and the infinitely wide parallel plate line). A longitudinal magnetic field will have associated with it transverse currents on the conductors. Since these transverse currents arise only because of the perturbation of the TEM mode into a mode with axial field components, they are small in magnitude compared with the axial current. Thus, the losses (proportional to current density squared) associated with the transverse currents will also be small compared with the losses due to the axial current. For a first approximation, the transverse currents may be neglected. When this is done, a variational integral for the propagation constant of a general lossy transmission line may be readily obtained.

The usual power-loss method of evaluating the attenuation constant of a lossy transmission line is based on an evaluation of the Joule heating loss in the conductors by assuming a current distribution identical with that for the loss-free line.[1] If P_L is the power loss per meter computed on this basis and P is the power flow along the line, the attenuation constant α is

$$\alpha = P_L/2P. \tag{1}$$

The variational integral to be presented will provide a justification of this method for general lossy transmission lines.

A Variational Formulation

For simplicity, a general two-conductor transmission line as in Fig. 1 will be considered although the

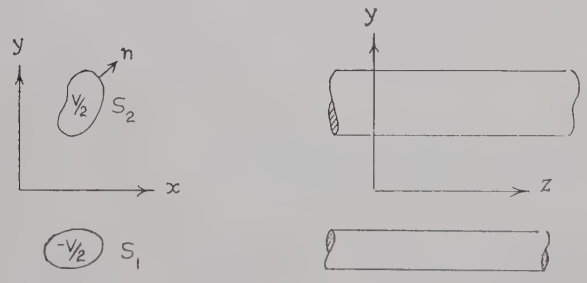


Fig. 1—A general two-conductor transmission line.

analysis is readily extended to multiconductor lines. If the current is assumed to be entirely in the axial direction and the medium surrounding the conductors to be homogeneous and isotropic, the field may be derived from a vector potential $\boldsymbol{A}$ having a z component $\psi(x, y)e^{-\gamma z}$ only. The relevant equations are

$$\boldsymbol{B} = \nabla \mathrm{x} \boldsymbol{A} \tag{2a}$$

$$\boldsymbol{E} = -j\omega \boldsymbol{A} + \frac{\nabla \nabla \cdot \boldsymbol{A}}{j\omega\epsilon\mu_0} \tag{2b}$$

$$\nabla^2 \boldsymbol{A} + k^2 \boldsymbol{A} = 0 \tag{2c}$$

[1] S. Ramo and J. R. Whinnery, "Fields and Waves in Modern Radio," John Wiley and Sons, Inc., New York, N. Y., 2nd ed., sec. 8.05: 1953.

where $k^2 = \omega^2\mu_0\epsilon$. For the present ϵ is assumed real. For a lossy dielectric medium surrounding the conductors ϵ is complex and shunt conductance losses are introduced. This requires only a trivial modification of the analysis, *i.e.*, replacing a real ϵ by a complex ϵ.

From (2b) and letting $\boldsymbol{A} = \boldsymbol{a}_z \psi e^{-\gamma z}$, it is found that

$$E_z = \frac{k_c^2}{j\omega\mu_0\epsilon} \psi e^{-\gamma z} \tag{3}$$

where $k_c^2 = k^2 + \gamma^2$. Provided the radius of curvature of the conductors S_1 and S_2 is much greater everywhere than the skin depth δ, the conductors will exhibit a surface impedance Z_m given by

$$Z_m = R_m + jX_m = (1 + j)/\sigma\delta \tag{4}$$

where σ is the conductivity and the skin depth δ is given by

$$\delta = (2/\omega\mu_0\sigma)^{1/2} \tag{5}$$

for a nonferrous material. The axial current density J is given by

$$\begin{aligned} J &= \boldsymbol{n}\mathrm{x}\boldsymbol{H} = \mu_0^{-1}\boldsymbol{n}\mathrm{x}(\nabla \mathrm{x} \boldsymbol{A}) \\ &= \mu_0^{-1}[\nabla(\boldsymbol{n}\cdot\boldsymbol{A}) - (\boldsymbol{n}\cdot\nabla)\boldsymbol{A}] = -\mu_0^{-1}\frac{\partial \boldsymbol{A}}{\partial n} \end{aligned} \tag{6}$$

where $\boldsymbol{n}$ is the unit outward normal from the conductors and $\boldsymbol{n}\cdot\boldsymbol{A} = 0$ since $\boldsymbol{n}$ and $\boldsymbol{A}$ are perpendicular. At the conductor surface $E_z = JZ_m$ and hence from (3) and (6) the boundary conditions for ψ are found to be

$$f\frac{\partial \psi}{\partial n} + \psi = 0 \text{ on } S_1, S_2 \tag{7}$$

where, for convenience, $j\omega\epsilon Z_m/k_c^2$ is denoted by f.

Let $\boldsymbol{E}_t e^{-\gamma z}$ be the transverse electric field. From (2b)

$$\boldsymbol{E}_t = \frac{j\omega\gamma}{k^2} \nabla\psi. \tag{8}$$

If there are no transverse currents, $\boldsymbol{n}\mathrm{x}\boldsymbol{E}_t$ or $\boldsymbol{n}\mathrm{x}\nabla\psi$ must vanish on S_1, S_2. This condition is in general incompatible with (7) except for the loss-free transmission line or lines with a high degree of symmetry, *e.g.*, coaxial line. For later use the solution for the loss-free line will be outlined. The vector potential z component will be taken as ψ_0. The propagation constant γ is equal to jk and $k_c = 0$. If the potentials of S_2 and S_1 are $V/2$ and $-V/2$ respectively, then

$$\begin{aligned} \psi_0 &= (\mu_0\epsilon)^{1/2}V/2 \text{ on } S_2 \\ &= -(\mu_0\epsilon)^{1/2}V/2 \text{ on } S_1. \end{aligned} \tag{9}$$

For the general case a variational integral for k_c^2 may be derived as follows. The equation satisfied by ψ is

$$\nabla_t^2\psi + k_c^2\psi = 0 \tag{10}$$

where ∇_t is the transverse operator $\mathbf{a}_x(\partial/\partial x)+\mathbf{a}_y(\partial/\partial y)$. Multiply (10) by ϕ, as yet an arbitrary function of x, y, that is regular at infinity, and integrate over the whole xy plane to get

$$k_c^2 \iint \phi\psi da = -\iint \phi\nabla_t^2\psi da.$$

Using Green's second theorem this may be rewritten as

$$k_c^2 \iint \phi\psi da = -\iint \psi\nabla_t^2\phi da + \oint_C \phi\frac{\partial\psi}{\partial n} dl - \oint_C \psi\frac{\partial\phi}{\partial n} dl \tag{11}$$

where n is the inward normal to the contour C. The contour C consists of the boundaries C_1, C_2 of the two conductors, the circle C_0 at infinity, and suitable cuts to make the region under consideration simply connected as in Fig. 2. Using the boundary condition (7) in (11)

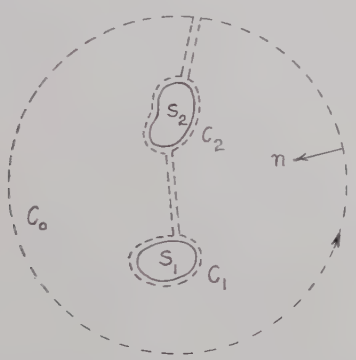


Fig. 2—Illustration of closed contour $C=C_0+C_1+C_2$.

gives

$$k_c^2 \iint \phi\psi da = -\iint \psi\nabla_t^2\phi da + \oint_{C_1+C_2}\left(\phi\frac{\partial\psi}{\partial n}+f\frac{\partial\psi}{\partial n}\frac{\partial\phi}{\partial n}\right)dl \tag{12}$$

since the contour integral around C_0 vanishes because ψ and ϕ are both regular at infinity. The first variation of (12) is now computed by standard methods to yield[2]

$$2k_c\delta k_c \iint \phi\psi da = -k_c^2 \iint \psi\delta\phi da - \iint \psi\nabla_t^2\delta\phi da - \iint \delta\psi(k_c^2\phi+\nabla_t^2\phi)da + \oint_{C_1+C_2}\left(\phi+f\frac{\partial\phi}{\partial n}\right)\frac{\partial\delta\psi}{\partial n}dl + \oint_{C_1+C_2}\left(\delta\phi+f\frac{\partial\delta\phi}{\partial n}\right)\frac{\partial\psi}{\partial n}dl.$$

The second term on the right may be transformed using Green's second theorem, *i.e.*,

$$-\iint \psi\nabla_t^2\delta\phi da = -\iint \delta\phi\nabla_t^2\psi da + \oint_{C_1+C_2}\left(\psi\frac{\partial\delta\phi}{\partial n}-\delta\phi\frac{\partial\psi}{\partial n}\right)dl$$

and hence, the variation in k_c is given by

$$2k_c\delta k_c \iint \phi\psi da = -\iint \delta\psi(\nabla_t^2\phi+k_c^2\phi)da - \iint \delta\phi(\nabla_t^2\psi+k_c^2\psi)da + \oint_{C_1+C_2}\left(\phi+f\frac{\partial\phi}{\partial n}\right)\frac{\partial\delta\psi}{\partial n}dl + \oint_{C_1+C_2}\left(\psi+f\frac{\partial\psi}{\partial n}\right)\frac{\partial\delta\phi}{\partial n}dl. \tag{13}$$

An examination of this result shows that the variation in k_c vanishes provided both ψ and ϕ satisfy the scalar Helmholtz equation (10) and the boundary condition (7). It now follows that if ϕ is replaced by ψ in (12) the resulting equation is a variational expression for k_c^2. Thus[3]

$$k_c^2 = \frac{-\iint \psi\nabla_t^2\psi da + \oint_{C_1+C_2}\left(f\frac{\partial\psi}{\partial n}+\psi\right)\frac{\partial\psi}{\partial n}dl}{\iint \psi^2 da}. \tag{14}$$

Substitution of a first-order approximate solution for ψ into (14) yields a solution for k_c^2 correct to the second order. For low-loss lines a suitable trial function to use in (14) is the corresponding solution ψ_0 for the loss-free line. The integral in the denominator serves as a normalization integral only. Once k_c^2 has been found, the propagation constant γ may be obtained at once from the relation $\gamma=(k_c^2-k^2)^{1/2}$.

Evaluation of k_c^2

The true potential function ψ_0 for the loss-free transmission line will be used as a trial function in (14). This function is a solution of Laplace's equation in the xy plane and hence $\nabla_t^2\psi_0=0$. Therefore, only the contour integral in (14) needs to be evaluated. Furthermore, ψ_0 will be normalized so that

$$\iint \psi_0^2 da = \mu_0\epsilon. \tag{15}$$

[2] P. M. Morse and H. Feshback, "Methods of Theoretical Physics," McGraw-Hill Book Co., Inc., New York, N. Y., vol. 2, sec. 9.4; 1953.

[3] This is not the only possible variational expression for k_c^2. However, this particular choice is a convenient one from which to obtain approximate results.

On S_1 and S_2 the current density J is equal to $-\mu_0^{-1}(\partial\psi/\partial n)$ with $\partial\psi/\partial n$ being negative on S_2 and positive on S_1. Also on S_1, $\psi_0 = -(\mu_0\epsilon)^{1/2}\ V/2$, and on S_2, $\psi_0 = (\mu_0\epsilon)^{1/2}\ V/2$. Substituting into (14) gives

$$k_c^2 = (\mu_0\epsilon)^{-1}\oint_{C_1+C_2}\left[\frac{j\omega\epsilon\mu_0 Z_m}{k_c^2}\mu_0 J^2 - \mu_0\psi J\right]dl$$

$$= \oint_{C_1+C_2}\frac{j\omega Z_m}{k_c^2}\mu_0 J^2 dl - Z_0\frac{V}{2}\oint_{C_2} J dl + Z_0\frac{V}{2}\oint_{C_1} J dl \tag{16}$$

where $Z_0 = (\mu_0/\epsilon)^{1/2}$. On S_1 the current is oppositely directed to that on S_2 so the last two integrals are together equal to $-Z_0 VI = -2Z_0 P$ where V is the potential difference between S_1 and S_2 and I is the total current flowing on one conductor and P is the power flow along the line. In the first integral on the right-hand side,

$$R_m\oint_{C_1+C_2} J^2 dl = 2P_L, \tag{17}$$

where P_L is the conductor loss per meter. Since the conductivity is finite, the magnetic field penetrates into the conductor and a net amount of magnetic energy is stored in this internal field. At the surface, the magnetic field is equal to the current density J in magnitude and decays exponentially with distance u into the conductor according to $e^{-u/\delta}$. Hence, the internal magnetic energy is

$$W_{mi} = \frac{1}{4}\mu_0\int_0^{u_0}\oint_{C_1+C_2} J^2 e^{-2u/\delta} du dl$$

where the integral over u need be taken only to some interior point u_0 where the field is negligible. Integrating over u gives

$$W_{mi} = \frac{1}{8}\mu_0\delta\oint_{C_1+C_2} J^2 dl = P_L/2\omega. \tag{18}$$

An internal inductance per meter may be defined by the relation

$$\tfrac{1}{4}I^2 L_i = W_{mi}$$

and from (18)

$$\omega L_i = R = 2P_L/I^2 \tag{19}$$

where R is the equivalent series resistance of the line per meter. Using these results the first integral in (16) becomes $(2kZ_0/k_c^2)(jP_L - 2\omega W_{mi})$. Finally, the equation for k_c^2 becomes

$$k_c^4 + 2Z_0 P k_c^2 - 2kZ_0(jP_L - 2\omega W_{mi}) = 0. \tag{20}$$

The solution for k_c^2 is

$$k_c^2 = -Z_0 P + [(Z_0 P)^2 + 2kZ_0(jP_L - 2\omega W_{mi})]^{1/2}$$

$$\approx \frac{k}{P}(jP_L - 2\omega W_{mi}) \tag{21}$$

where the binomial expansion could be used since P_L and W_{mi} are small compared with P for low-loss lines. Replacing k_c^2 by $k^2 + (j\beta + \alpha)^2$ gives

$$2\beta\alpha = \frac{kP_L}{P} \tag{22a}$$

$$k^2 - \beta^2 + \alpha^2 = -\frac{2W_{mi}}{P}\omega k. \tag{22b}$$

Since the losses are assumed small $\beta \approx k$ and $k^2 - \beta^2$ may be replaced by $2k(k-\beta)$. The right-hand side of (22b) is of magnitude $2\alpha k$ and, hence, α^2 is negligible in comparison. Therefore, (22) reduces essentially to the more familiar expressions

$$\alpha = P_L/2P \tag{23a}$$

$$\beta = k + \omega W_{mi}/P. \tag{23b}$$

For a lossless line $k = \omega\sqrt{LC}$ where L and C are the inductance and capacitance per unit length. Also, the characteristic impedance Z_c is equal to $(L/C)^{1/2}$ and

$$P_L = \tfrac{1}{2}RI^2, \qquad P = \tfrac{1}{2}Z_c I^2.$$

Therefore, (23) may be rewritten as follows:

$$\alpha = R/2Z_c \tag{24a}$$

$$\beta = \omega[C(L + L_i)]^{1/2}. \tag{24b}$$

The phase constant β is increased by an amount corresponding to the increase in the inductance of the line per meter due to addition of internal inductance. Since ωL_i and R are introduced together because of the finite conductivity, the series resistance of the line may be considered as equivalent to making the permeability of the medium surrounding the conductors complex. This provides a formal analogy between series resistance loss and shunt conductance loss as follows:

Lossless Line	*Lossy Line*
$\epsilon = \epsilon$	$\epsilon = \epsilon[1 - jG/\omega C]$
$\mu = \mu_0$	$\mu = \mu_0\left[1 + \frac{L_i}{L} - \frac{jR}{\omega L}\right]$

where G is the shunt conductance, L_i the internal inductance, L the external inductance, and R the series resistance of the line.

Conclusion

The use of a variational integral for k_c^2 verifies the validity of the usual power-loss method for evaluating the attenuation constant of low-loss transmission lines. In addition, it provides a method whereby k_c^2 can be evaluated to any desired degree of accuracy by using a trial function containing several variational parameters and determining these so as to make (14) stationary. The method is, however, limited to those cases where the transverse currents can be neglected in comparison with the axial currents.

Measurement of Bandwidth of Microwave Resonator by Phase Shift of Signal Modulation*

D. S. LERNER† AND H. A. WHEELER†

Summary—Bandwidth is measured by transmission of a signal with sine-wave modulation through a microwave resonator under test. The modulation frequency is adjusted so that the envelope is delayed 45° with respect to the input, indicating that the two sideband frequencies are separated by the half-power bandwidth. The resonance ratio (Q) is then equal to the ratio of carrier frequency over twice the modulation frequency. This depends on observations of these frequencies and the modulation phase shift, but not on the amplitude. It is insensitive to detuning or incidental frequency variation of the resonator or the signal. In a resonant cavity tested, an observed bandwidth of 30 kc at 700 mc indicated that $Q=23{,}300$.

THIS PAPER describes a technique for measuring the bandwidth of a high-Q resonator by observing the transmission phase shift of the envelope of a modulated signal. Furthermore, this technique is suited for microwave distributed-circuit measurements as distinguished from techniques which are suited only for lumped-circuit measurements at lower frequencies. In general, prior methods have depended on amplitude observations, whereas the subject method does not.

Microwave resonators typically have a high Q, much greater than 1000. They are utilized in narrow-band filters, wavemeter cavities, and echo boxes. The common method [1], [2] of measuring the bandwidth would require an observation of the resonance curve, noting the frequency difference between the half-power (3-db) points on the sides. Much ingenuity has been applied to this method, such as sweeping the curve, much expanded, on a scope and then superimposing amplitude and frequency markers for ease and precision of observation.

A special method of microwave measurement, commonly applied to the echo box, relies on observing the envelope of the damped free oscillation following excitation by a short pulse. This also relies on observation of amplitude, particularly its rate of decay with time.

In a lumped resonant circuit, the "Q meter" relies on directly observing the ratio of the voltages across one reactance and the net resistance [3], [4]. It relies on amplitude measurement also, and, in addition, it is not generally applicable to distributed-circuit resonators.

As an introduction to the principle of the subject method, Fig. 1 shows the well-known behavior of transmission through a resonator at frequencies near resonance. The upper curve (a) shows the amplitude peak at f_0 and the two side points separated by the half-power (3-db) bandwidth Δf. The lower curve (b) shows the phase slope and the corresponding two side points at $\pm 45°$ relative to the center.

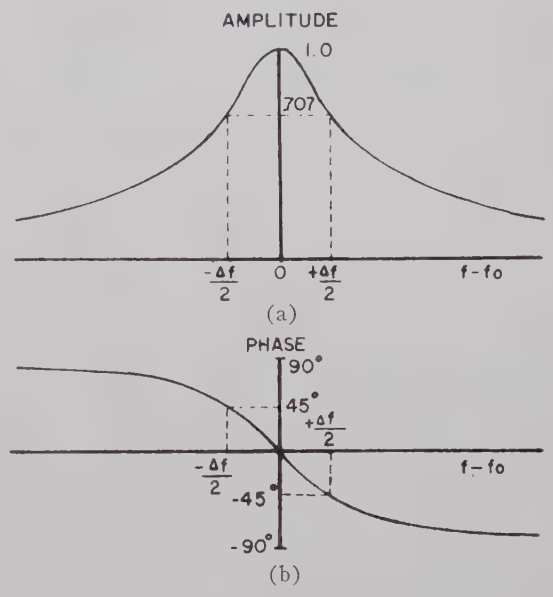


Fig. 1—Transmission characteristics of single-tuned resonator.

If a carrier with sine-wave amplitude modulation is transmitted through such a resonator, with the carrier at the resonator center frequency and the sidebands at the half-power frequencies, the modulation envelope is delayed by 45° at the output. This phase lag with respect to the input results from the 45° phase shift at each of the sideband frequencies relative to that at the carrier frequency.

The block diagram of Fig. 2 shows the means of utilizing this principle for determining bandwidth. The signal generator provides an RF carrier with sidebands to give a sine-wave envelope of amplitude modulation. This is transmitted through the RF resonator under test. The input and output signals of the resonator are rectified in a detector, and the resulting modulation waves are amplified. The two resulting sine waves are applied to a phase meter to indicate their phase difference. An auxiliary amplitude indicator is provided for assuring that the amplitude is within the wide limits of normal operation of the phase meter.

In operation, the signal is tuned to the resonator center frequency and the modulation frequency is adjusted to give an indication of 45° on the phase meter. The details of this procedure will be given after a further description of the equipment.

* Manuscript received by the PGMTT, December 10, 1959; revised manuscript received, December 28, 1959. Presented at the PGMTT Annual Symposium, Cambridge, Mass., June 2, 1959.

† Wheeler Labs. Inc., Smithtown, N. Y.

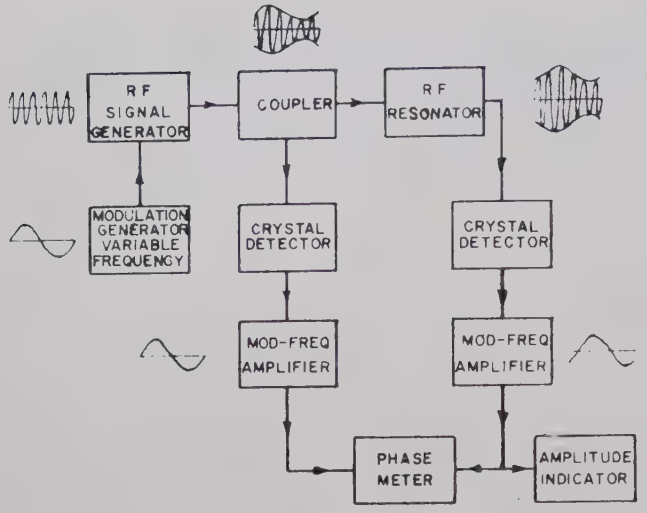


Fig. 2—Block diagram of measurement equipment.

In the signal generator, the carrier frequency must be tunable to the resonator frequency, and should be stable within about 1/10 of the bandwidth being measured. The signal generator must have provision for sine-wave modulation. This is readily accomplished internally in the types of circuits commonly used in signal generators below 1000 mc. At higher frequencies, where internal linear amplitude modulation may not be available, a crystal or magnetic modulator may be inserted in the output of the signal generator. The modulation factor should be kept moderately low (say less than 0.3) so as to hold down any envelope distortion in modulation, which may be a cause of error.

The modulation generator must provide a sine wave of a frequency which will be made equal to one-half the bandwidth. A continuously adjustable frequency is provided because the setting of the modulation frequency is part of the measurement procedure.

The coupler provides a reference signal by sampling the input to the cavity. Its coupling loss should approximately equal the insertion loss of the cavity. If the coupler is not directional, the line from coupler to resonator should be padded so that the reflections from the resonator will not cause errors.

The detectors and amplifiers are similar, and must have the same phase shift at the modulation frequency. The detectors may have any law of rectification that is not too unusual. The measurement procedure corrects for errors caused by distortion in detection, provided that the distortion is not so large as to change the number or the sequence of the zero crossings. The modulation-frequency amplifiers are needed to provide a sufficiently high amplitude for the operation of the phase meter.

The phase meter may be of any (within limits) of the several commercial types which indicate relative phase angle independent of amplitude and frequency [5]. Such a meter may rely on the zero crossings of the two applied signals, as will be assumed here.

The amplitude indicator is an oscilloscope or a voltmeter which indicates the output from the resonator. It is required merely to indicate whether the amplitude is within the operating limits of the phase meter.

In operation, the output of the RF signal generator is applied to the resonator under test. A sample of the applied signal is coupled to the first crystal detector, and serves as a reference. The applied signal is transmitted through the cavity and thereby has its modulation envelope changed. The modulation factor is reduced and the phase angle of the modulation is retarded; only the latter change is used for the measurement. The transmitted signal is detected and amplified like the reference signal. Both signals are applied to the phase meter, which then indicates directly their phase difference. The amplitude indicator may be used for setting the initial tuning in addition to monitoring the signal level to insure that it is within the range specified for the phase meter.

The measurement procedure is as follows:

a) The modulation frequency is set at some value less than the bandwidth to be measured, so that the detected modulation can be used for setting the carrier frequency.
b) The carrier frequency is set by adjusting the RF signal generator for maximum transmission through the resonator, with the aid of the amplitude indicator. This operation may be made completely independent of amplitude by tuning the RF signal generator for maximum phase shift instead of maximum amplitude. These alternative procedures yield the same result.
c) The frequency of the modulation generator is adjusted for 45° phase shift as indicated by the phase meter. The effects of any detector distortion can be corrected by making two measurements, as discussed below.
d) The bandwidth is equal to the frequency difference between sidebands, or twice the modulation frequency. The resonance ratio (Q) is therefore the carrier frequency divided by twice the modulation frequency.

The wave forms involved in the phase measurement are shown in Fig. 3. The like sine waves (a) illustrate the modulation signals applied to the phase meter in the absence of distortion. The time intervals marked 45° indicate the intervals used for the phase indication. The unlike wave forms (b) represent the outputs of detectors having different laws of rectification. The distortion does not change the time of positive and negative peaks of the signal, but it does change the time of the zero crossing. For a sine-wave modulation envelope, this distortion is symmetrical about the peaks. That is, if one zero crossing is advanced, the other is retarded by an equal amount. This is illustrated by the two time intervals which deviate from 45° by $\pm e$. By using the average of two phase measurements, using respectively

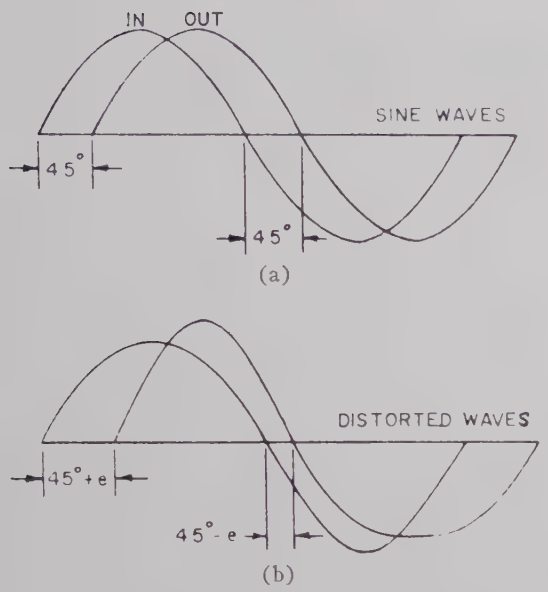


Fig. 3—Detected waveforms of modulation envelope.

the positive-going and the negative-going zero crossings, the effects of a moderate amount of distortion may be cancelled. For the particular phase meter used, this is accomplished by interchanging the two inputs. In general, the two measurements may be made by reversing the polarity of both inputs.

The bandwidth is indicated directly as twice the modulation frequency. This is inherently more accurate than separately measuring the half-power frequencies and then subtracting these frequencies to evaluate their small difference. Amplitude variations of the signal generator have no effect, since the measurement utilizes phase angle only. The envelope phase shift through the resonator has its greatest value with the carrier at the resonant frequency. Therefore, deviations from this frequency have only a second-order effect on phase shift, and the measurement is relatively insensitive to frequency drift of the signal generator and the resonator under test. A small amount of frequency modulation at the amplitude modulation frequency does not impair the accuracy, because its sidebands do not affect the detected amplitude.

For example, measurements by this procedure have been used to determine a bandwidth of 30 kc at a frequency of 700 mc, which gives a Q of 23,300. The accuracy is estimated to be within 3 per cent, corresponding to an error of 0.7° in measuring the phase. If individual half-power frequencies were measured to obtain this result, they would have to be determined within one part in 2 million. As compared with other techniques that are available for determining the bandwidths of high-Q resonators, the modulation phase method offers advantages of simplicity in operation and accuracy of determination.

Bibliography

[1] M. Wind and H. Rapaport, "Handbook of Microwave Measurements," Microwave Res. Inst., Polytechnic Inst. of Brooklyn, Brooklyn, N. Y.; 1954. (Section V, "Measurement of Q," mentions phase shift of 45° at half-power frequencies.)

[2] C. G. Montgomery, "Technique of Microwave Measurements," Rad. Lab. Ser., McGraw-Hill Book Co., Inc., N. Y., chs. 5 and 6; 1947.

[3] F. E. Terman, "Radio Engineers' Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., p. 916; 1943.

[4] "Instructions and Manual of RF Measurements for the Q-Meter Type 100A," Boonton Radio Corp., Boonton, N. J.; March, 1935. (Talk by C. J. Franks and W. D. Loughlin at Rochester Fall Meeting, Rochester, N. Y.; November, 1934.)

[5] Y. P. Yu, "Coincident slicer measures phase directly," *Electronics*, vol. 31, pp. 99–101; September 12, 1958. (Description of phase meter.)

[6] D. C. King, "Measurements at Centimeter Wavelength," D. Van Nostrand Co., Inc., New York, N. Y.; 1952. (Resonant circuits, pp. 128–147.)

[7] E. L. Ginzton, "Microwave Measurements," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 9; 1957. (Measurement of Q, various methods.)

[8] F. H. James, "A method for the measurement of very high Q-factors of electromagnetic resonators," Proc. IRE, vol. 106, pp. 489–492; September, 1959. (Recent proposal of the subject method.)

A Y-Junction Strip-Line Circulator*

U. MILANO†, J. H. SAUNDERS†, AND L. DAVIS, JR.†

Summary—The theoretical approach to the three-port symmetrical circulator is reviewed and presented in a form valid for the most general waveguide case.

A strip-line *Y*-junction circulator is described and the performance of different units in the band 800–1600 mc is illustrated.

The new type of device described offers, for the low-frequency region of the microwave spectrum, advantages of simple design, light weight, and great compactness with respect to the classical types. When operated with a permanent magnet it gives—in a bandwidth of about 4 per cent—isolation greater than 20 db, insertion loss $\leq$0.4 db, and input VSWR $\leq$1.20.

I. Introduction

RECENTLY, several people have worked on microwave circulators consisting of symmetrical waveguide junctions containing symmetrically located pieces of ferrite.[1–5]

The experimental results which have been published concern, in general, turnstile junctions or *n*-port star junctions of rectangular waveguides; the present paper deals with a TEM structure realized in strip-line which can offer distinct advantages of compactness, small weight, and structural simplicity when applications in the low-frequency region of the microwave spectrum are considered.

First of all, some theoretical considerations are reviewed which can illustrate the behavior of this type of circulator, and afterwards some experimental results obtained in the range 800–1600 mc are presented.

II. The Mathematical Approach to the Three-Port Symmetrical Circulator

The scattering matrix formalism and the application of group theory have been proven to be very useful for approaching the theoretical problems of symmetrical waveguide circulators.

Treuhaft[6] has shown the correspondence between the scattering matrix of circulators and the cyclic substitutions of the group theory, and he has investigated on this basis some general symmetry properties of the circulators. Auld[4] has applied these mathematical tools to the general problem of the synthesis of symmetrical waveguide circulators; Sirvetz[7] has applied the scattering matrix formalism for a perturbation treatment of the *Y*-junction circulator.

One of the interesting features of the scattering matrix approach is that it can be applied to the most general type of waveguide, including in the term "waveguide" also the TEM structures. Therefore, we will follow this approach and we will consider in this section the most general case of a lossless symmetrical *Y*-junction of three microwave lines containing at its center a ferrite medium located in such a way that it does not alter the symmetry of the junction (Fig. 1) and magnetized along the axis of the junction.

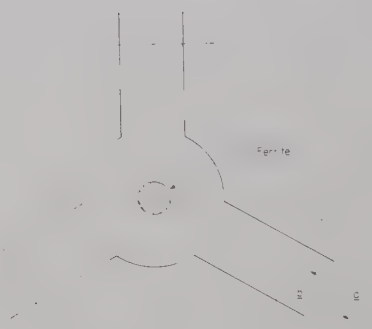


Fig. 1—Microwave *Y*-junction.

In order that this junction be a circulator—according to Treuhaft's definition[6]—its scattering matrix must operate on the incident waves so as to produce the same result as a cyclic substitution.[8] That is to say, if we indicate with E_1^i, E_2^i, E_3^i, the transverse components of the incident electric fields and with E_1^r, E_2^r, E_3^r, the transverse components of the reflected electric fields at the reference planes of the junction, the scattering matrix must produce the same result as the substitution

$$[E_1^r\ E_2^r\ E_3^r] = \{(1\ 2\ 3) \rightarrow [E_1^i\ E_2^i\ E_3^i]\}$$
$$= [E_2^i\ E_3^i\ E_1^i] \qquad (1)$$

* Manuscript received by the PGMTT, November 13, 1959; revised manuscript received, January 26, 1960. This paper was presented at the PGMTT Natl. Symp., Cambridge, Mass., June 1–3, 1959.

† Research Div., Raytheon Company, Waltham, Mass.

[1] P. J. Allen, "The turnstile circulator," IRE Trans. on Microwave Theory and Techniques, vol. MTT-4, pp. 223–227; October, 1956.

[2] T. Schaug-Pettersen, ONR London Tech. Rept. ONRL 111-57; September, 1957.

[3] H. N. Chait and T. E. Curry, "*Y* circulator," *J. Appl. Phys.*, suppl. to vol. 30, pp. 152S–153S; April, 1959.

[4] B. A. Auld, "The synthesis of symmetrical waveguide circulators," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, pp. 238–246; April, 1959.

[5] S. Yoshida, "*X* circulator," Proc. IRE, vol. 47, p. 1150; June, 1959.

[6] M. A. Treuhaft, "Network properties of circulators based on the scattering concept," Proc. IRE, vol. 44, pp. 1394–1402; October, 1956.

[7] M. H. Sirvetz, "The *Y*-Junction Microwave Circulator," Raytheon Co., Waltham, Mass., Tech. Memo T-143; March, 1959.

[8] By the cyclic substitution (*abc*) operating on a sequence of three elements, we mean an operation that substitutes the element *a* by *b*, *b* by *c*, and *c* by *a*. If, for example, we apply such a substitution to the sequence *bac* and we indicate it with the symbol $\{(abc) \rightarrow bac\}$, we obtain the result:

$$\{(abc) \rightarrow bac\} = cba.$$

for one sense of circulation, or the substitution

$$[E_1^r\ E_2^r\ E_3^r] = \{(1\ 3\ 2) \rightarrow [E_1^i\ E_2^i\ E_3^i]\} = [E_3^i\ E_1^i\ E_2^i] \tag{2}$$

for the opposite sense of circulation.

These substitutions must correspond, using the matrix notation, to

$$\begin{bmatrix} E_1^r \\ E_2^r \\ E_3^r \end{bmatrix} = \mathbf{S}^{(c)} \begin{bmatrix} E_1^i \\ E_2^i \\ E_3^i \end{bmatrix} \tag{3}$$

where $\mathbf{S}^{(c)}$ indicates the scattering matrix of the three-port circulator. We will restrict ourselves to consider only the sense of circulation corresponding to the substitution (1), and one immediately sees that the scattering matrix required to perform this operation is

$$\mathbf{S}^{(c)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \tag{4}$$

The form (4) assumes a suitable choice of the reference planes to correspond to a zero phase shift through the junction.

We now represent with the complex number a_i ($i=1, 2, 3$) the amplitude and phase of the transverse electric field of the incident wave at the ith reference plane, normalized in such a way that the average incident power is given by $\frac{1}{2}a_i^* a_i$; we also indicate with b_i the corresponding measure of the emergent wave. If we indicate with $\mathbf{a}$ and $\mathbf{b}$ the column vectors

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

we can write (3) in the form

$$\mathbf{b} = \mathbf{S}^{(c)}\mathbf{a}. \tag{3'}$$

Let us consider the eigenvalue equation of the square matrix $\mathbf{S}^{(c)}$:

$$\mathbf{S}^{(c)}\mathbf{a}^{(c)} = s^{(c)}\mathbf{a}^{(c)} \tag{5}$$

where $\mathbf{a}^{(c)}$ is now an eigenvector and $s^{(c)}$ an eigenvalue.

Eq. (5) has a nonvanishing solution for $\mathbf{a}^{(c)}$ if the following condition is satisfied:

$$\det \left| \mathbf{S}^{(c)} - s^{(c)}\mathbf{I} \right| = 0 \tag{6}$$

where $\mathbf{I}$ indicates the unity matrix.

The solution of the characteristic equation obtained by (6) gives the following set of nondegenerate eigenvalues:

$$\begin{aligned} s_1^{(c)} &= 1 \\ s_2^{(c)} &= e^{j(2\pi/3)} \\ s_3^{(c)} &= e^{j(2\pi/3)}. \end{aligned} \tag{7}$$

By substituting these values in (5), one obtains for the normalized eigenvectors the set

$$\mathbf{a}^{(1)} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \quad \mathbf{a}^{(2)} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ e^{-j(2\pi/3)} \\ e^{j(2\pi/3)} \end{bmatrix};$$

$$\mathbf{a}^{(3)} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ e^{j(2\pi/3)} \\ e^{-j(2\pi/3)} \end{bmatrix}. \tag{8}$$

It is rather interesting to compare these results with those one obtains in the case of the reciprocal Y-junction. If we indicate with $\mathbf{S}^{(r)}$ the matrix of the reciprocal junction and with $S_{ij}^{(r)}$ the generic element of the matrix, we realize, by inspecting the symmetry properties of the junction, that the following equalities must hold:

$$\begin{aligned} S_{11}^{(r)} &= S_{22}^{(r)} = S_{33}^{(r)} = S' \\ S_{12}^{(r)} &= S_{21}^{(r)} = S_{13}^{(r)} = S_{31}^{(r)} = S_{23}^{(r)} \\ &= S_{32}^{(r)} = S''. \end{aligned} \tag{9}$$

That is to say,

$$\mathbf{S}^{(r)} = \begin{bmatrix} S' & S'' & S'' \\ S'' & S' & S'' \\ S'' & S'' & S' \end{bmatrix}. \tag{10}$$

By solving the determinental equation

$$\det \left| \mathbf{S}^{(r)} - s^{(r)}\mathbf{I} \right| = 0, \tag{11}$$

one finds for the eigenvalues of $\mathbf{S}^{(r)}$

$$\begin{aligned} s_1^{(r)} &= S' + 2S'' \\ s_2^{(r)} &= s_3^{(r)} = S' - S''. \end{aligned} \tag{12}$$

The two-fold degeneracy of the eigenvalues $s_2^{(r)}$ and $s_3^{(r)}$ leads to an ambiguity in the definition of the eigenvectors $\mathbf{a}_2^{(r)}$ and $\mathbf{a}_3^{(r)}$; for our purposes, however, it is interesting to notice that one possible set of eigenvectors of $\mathbf{S}^{(r)}$, corresponding to the above indicated eigenvalues, is given by the set (8), as one can easily verify.

The eigensolutions corresponding to the set of eigenvectors (8) have been described in the literature.[9] For the eigensolution corresponding to the eigenvector $\mathbf{a}^{(1)}$, the electromagnetic field at the center of the junction has only components parallel to the axis of the junction. For the eigensolutions corresponding to the eigenvectors $\mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$, the axial fields vanish at the center of the junction, but the transverse components of the electric and magnetic field give rise there to circularly-polarized waves rotating in one sense for the eigenvector $\mathbf{a}^{(2)}$ and in the opposite sense for the eigenvector $\mathbf{a}^{(3)}$.

When we introduce into the reciprocal Y-junction a symmetrically-located piece of ferrite, magnetized along the axis of the junction, the relationship $S_{ik} = S_{ki}$ no

[9] C. G. Montgomery, R. H. Dicke, and E. M. Purcell, "Principles of Microwave Circuits," McGraw-Hill Book Co., Inc., New York, N. Y., chs. 5 and 12; 1948.

longer holds, and we obtain a scattering matrix $S^{(f)}$ that assumes the general form

$$S^{(f)} = \begin{bmatrix} S' & S'' & S''' \\ S''' & S' & S'' \\ S'' & S''' & S' \end{bmatrix}. \tag{13}$$

The eigenvector (8) of $S^{(r)}$ are, however, still eigenvectors of $S^{(f)}$ because the ferrite does not alter the symmetry of the junction either from the geometrical point of view or from the electromagnetic point of view.

In the device considered here, the ferrite sample, whose transverse dimensions we suppose small in comparison with the wavelength, is located in a region where the RF magnetic field is either directed along the axis of the junction for the first eigensolution, or is circularly-polarized for the other two eigensolutions. In each case, the ferrite presents an effective scalar permeability which affects the corresponding eigenvalues. In particular, $s_2^{(f)}$ and $s_3^{(f)}$ will assume different values because of the different effective permeabilities pertaining to the two corresponding eigensolutions. In other words, the introduction of the ferrite has the effect of removing what has been called the "reciprocity degeneracy" of the Y-junction.[10]

We now observe that $S^{(f)}$ can be diagonalized by a matrix X having for its columns the eigenvectors of $S^{(f)}$, *i.e.*, the set (8).[11]

By this we mean that for such a matrix X, the following relationship holds:

$$X^{-1}S^{(f)}X = \Lambda \tag{14}$$

where

$$\Lambda = \begin{bmatrix} s_1^{(f)} & 0 & 0 \\ 0 & s_2^{(f)} & 0 \\ 0 & 0 & s_3^{(f)} \end{bmatrix}. \tag{15}$$

In a different form, (14) can be written

$$S^{(f)} = X\Lambda X^{-1}; \tag{14'}$$

and by multiplying out the right side of (14'), we obtain

$$S_{11}^{(f)} = S_{22}^{(f)} = S_{33}^{(f)} = \frac{s_1^{(f)} + s_2^{(f)} + s_3^{(f)}}{3} = S' \tag{16}$$

$$S_{12}^{(f)} = S_{23}^{(f)} = S_{31}^{(f)} = \frac{s_1^{(f)} + s_2^{(f)}e^{j(2\pi/3)} + s_3^{(f)}e^{-j(2\pi/3)}}{3} = S'' \tag{17}$$

$$S_{13}^{(f)} = S_{32}^{(f)} = S_{21}^{(f)} = \frac{s_1^{(f)} + s_2^{(f)}e^{-j(2\pi/3)} + s_3^{(f)}e^{j(2\pi/3)}}{3} = S'''. \tag{18}$$

In order that the junction of Fig. 1—whose matrix we have called $S^{(f)}$—be a circulator, it is necessary that the eigenvalues $s^{(f)}$ coincide with the eigenvalues $s^{(c)}$ of $S^{(c)}$; that is to say, if we represent them on a complex plane, the extremes of the vectors representing them must lie on a circle of unit radius and they must be equally spaced at angles of 120°.

In this situation, the right side of (16) becomes zero, which simply says that, in order that the junction of Fig. 1 be a circulator, it must appear matched at any port. This fact is included in the circulator definition we have assumed, but it is worthwhile to notice, from a practical point of view, that the condition $S'=0$ not only is necessary for the circulator action of a lossless Y-junction, but it is also sufficient, as one can easily see by imposing the unitary condition on the matrix $S^{(f)}$, together with $S'=0$.[12] Therefore, one can utilize this property for the experimental adjustment of the circulator as we will see in the next section.

III. The Experimental Adjustment of the Y-Junction Strip-Line Circulator

The previous discussion has essentially shown that, in order to obtain a circulator action from the junction schematically illustrated in Fig. 1, the nonreciprocal action of the ferrite has to be such as to remove the degeneracy of the eigenvalues s_2 and s_3 of $S^{(r)}$ and to obtain the coincidence of the new set of eigenvalues $s_1^{(f)}$, $s_2^{(f)}$, $s_3^{(f)}$ with the set (7) of the eigenvalues $s_i^{(c)}$.

As the vectors representing the $s^{(f)}$ can always be rotated through an arbitrary angle by properly choosing the reference planes of the junction, the above conditions can be expressed as follows:

$$s_2^{(f)} = s_1^{(f)}e^{-j(2\pi/3)} \tag{19}$$

$$s_3^{(f)} = s_1^{(f)}e^{j(2\pi/3)}. \tag{20}$$

This shows that the experimental adjustment of the circulator requires the adjustment of two phase angles and therefore the variation of two physical parameters.

In our case, the geometrical configuration of the circulator was realized according to the general structure of Fig. 2. Two disks of ferrite are placed at the center of the junction and they are magnetized along the direction perpendicular to their plane. All the results shown in the next figures refer to the case of disks of polycrystalline YIG polarized above resonance.

The adjustment of the circulator at a fixed frequency was generally done by using as variable parameters the diameter d of the disks and the value H_0 of the polarizing field. These two variable parameters were optimized until the junction appeared to be matched or nearly matched. In this condition, the circulator action is automatically insured.

For any fixed frequency, there is one optimum value of the diameter of the disks which achieves the matched

[10] D. M. Kerns, "Analysis of symmetrical waveguide junctions," *J. Res. NBS*, vol. 46, pp. 267–282; April, 1951.

[11] H. Margenau and G. M. Murphy, "The Mathematics of Physics and Chemistry," D. Van Nostrand Co., New York, N. Y., ch. 10, sec. 15; 1943.

[12] H. J. Carlin, "Principles of gyrator networks," *Proc. Symp. on Modern Advances in Microwave Techniques*, Polytechnic Inst. of Brooklyn, Brooklyn, N. Y., pp. 175–204; November, 1954.

Fig. 2—Y-junction strip-line circulator.

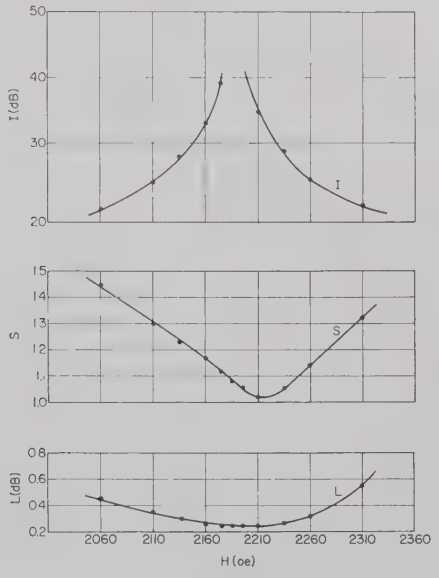


Fig. 3—Performance vs polarizing field at a fixed frequency; $f = 1030$ mc.

condition, and this d_{opt} is a continuously decreasing function of the frequency.

By keeping the mechanical dimensions of the junction unchanged and by varying the thickness of the disks, one has another variable parameter that also can be utilized for the adjustment of the circulator. At a fixed frequency and with a fixed mechanical structure, a reduction of the thickness of the disks leads to a greater value of d_{opt}.

Fig. 3 shows some typical results obtained at a fixed frequency by changing the magnetic field when $d \simeq d_{opt}$. We have indicated by I the attenuation between the two decoupled arms ("isolation"), with L the insertion loss between input and forward arm, and with S the VSWR seen at the input arm.

The isolation measurements have always been taken by using a load with an input VSWR of 1.02 on the forward arm and the measurements of the insertion loss have always been taken by using a tunable bolometer.

It appears from the data of Fig. 3 that it is possible, at a single frequency, to associate isolations greater than 30 db with insertion loss of 0.3 db or less, and input VSWR of 1.05.

In Fig. 3, the maximum value of isolation does not correspond to the minimum value of the input VSWR. This is believed to be due to small asymmetries present in the structure.

In Fig. 3 and in the following figures, the values of isolation do not coincide in general with the values one could calculate from the corresponding values of input VSWR in the hypothesis of a lossless circulator.

This may be due to asymmetry effects, but it is shown in the Appendix that the effect of small departures from the optimal conditions for H_0 and d combined with the effect of the losses can also account for this type of deviation.

By choosing the appropriate dimensions of the ferrite and the value of the polarizing magnetic field, it has been possible, by using the mechanical structure indicated in Fig. 2, to cover the range 800–1600 mc with a performance always comparable to that presented in Fig. 3.

The circulator can be operated with a small permanent magnet, and therefore one obtains a simple, very compact device which can give a useful bandwidth of about 4 per cent where the isolation is not less than 20 db, the insertion loss not greater than 0.4 db, and the input VSWR not greater than 1.20.

Figs. 4–6 show some examples of performance obtained in the regions of 800, 1000, and 1300 mc, with some laboratory units polarized with a fixed field.

IV. Conclusion

The mathematical approach to the three-port symmetrical circulator has been reviewed and it has been presented in a form valid for the most general waveguide case. The behavior of the rectangular waveguide three-port symmetrical circulator has been explained by other authors,[3] on the basis of a field displacement effect. The approach illustrated in this paper, which was first used by Auld for the general problem of the synthesis of symmetrical circulators, does not require the hypothesis of a field displacement effect and it seems to the authors more appropriate for the specific experimental type of circulator illustrated here.

The Y-junction strip-line circulator described in the present paper is a new type of device which represents an attractive solution for circulators operating in the low-frequency region of the microwave spectrum where it offers, in comparison with the classical types, distinct advantages of simple design, light weight, and great compactness. This application appears to be of particular interest for use in connection with masers and parametric amplifiers operating in the frequency region mentioned above.

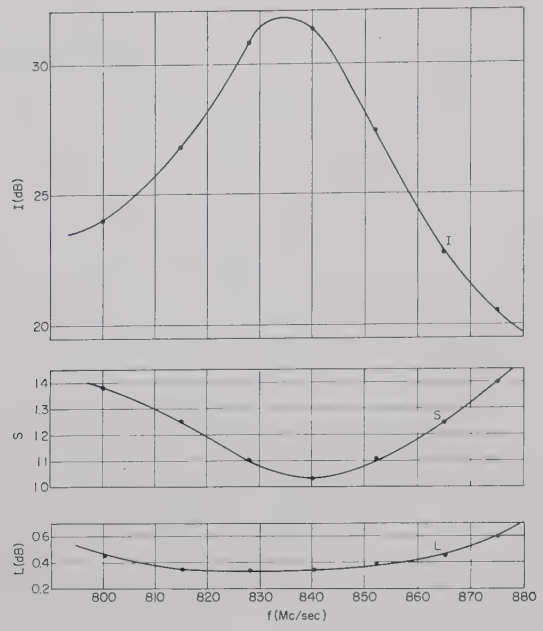


Fig. 4—Performance vs frequency for a fixed polarizing field in the region of 830 mc.

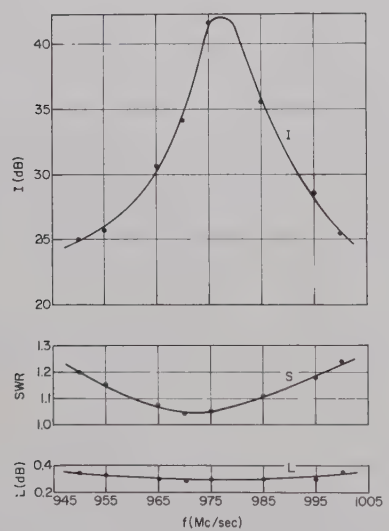


Fig. 5—Performance vs frequency for a fixed polarizing field in the region of 975 mc.

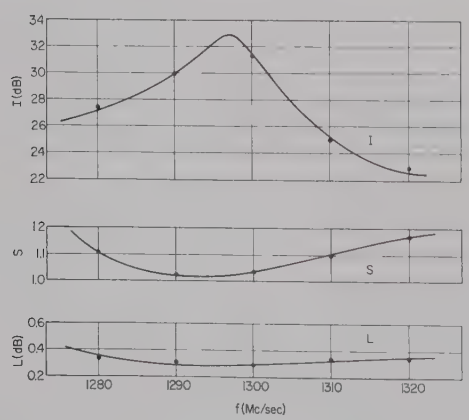


Fig. 6—Performance vs frequency for a fixed polarizing field in the region of 1300 mc.

Appendix

The Behavior of an Imperfect Y-Junction Circulator

In Section II, it was mentioned that when one introduces inside a lossless microwave Y-junction a symmetrically-located piece of ferrite, magnetized along the axis of the junction, one obtains a new nonreciprocal junction with the following characteristics:

1) The scattering matrix $\boldsymbol{S}^{(f)}$ of the nonreciprocal junction is given in the most general case by (13).
2) The eigenvectors (8) of the scattering matrix $\boldsymbol{S}^{(r)}$ of the reciprocal junction are still eigenvectors of $\boldsymbol{S}^{(f)}$.
3) The reciprocity degeneracy of the eigenvalues $s_2^{(r)}$ and $s_3^{(r)}$ is removed and $\boldsymbol{S}^{(f)}$ has three different eigenvalues $s_i^{(f)} = |s_i| e^{j\phi_i}$ $(i=1, 2, 3)$, ϕ_i being a real number.
4) The elements of the scattering matrix $\boldsymbol{S}^{(f)}$ are given, as functions of the eigenvalues $s_1^{(f)}$, $s_2^{(f)}$, $s_3^{(f)}$, by the formulas (16)–(18).

One can now have three different situations.

a) The junction containing ferrite is lossless, the $|s_i|$ are equal to unity, and the unit vectors representing the eigenvalues on the complex plane are equally spaced at angles of 120°

$$\phi_2 - \phi_3 = \phi_3 - \phi_1 = \frac{2\pi}{3}. \tag{21}$$

b) The junction has losses, but the conditions (21) are fulfilled.
c) The junction has losses and the conditions (21) are not fulfilled.

In this last case, we will call the junction an "imperfect" Y-junction circulator.

If the dimensions of the ferrite element and of the value of the polarizing magnetic field are not optimal, one does not satisfy conditions (21). In this case, one can write:

$$\phi_2 - \phi_1 = \frac{4\pi}{3} + \delta_2$$

$$\phi_3 - \phi_1 = \frac{2\pi}{3} + \delta_3. \tag{22}$$

As far as the influence of the losses is concerned, we can observe that the eigenvalues s_i $(i=1, 2, 3)$ represent reflection coefficients and, therefore, in the absence of losses $|s_i| = 1$. In the presence of losses, however, one will have $|s_i^{(f)}| \leqq 1$, and we can take this fact into account if we represent the eigenvalues of $\boldsymbol{S}^{(f)}$ as exponential functions having complex exponents. That is, in the case of an imperfect Y-junction circulator, the most general expression of the generic eigenvalue will be:

$$s_i = e^{(-\alpha_i + j\phi_i)} \quad (i = 1, 2, 3). \tag{23}$$

In particular, if, for a suitable choice of the reference planes of the junction, we assume $\phi_1 = 0$, we can write the set of eigenvalues as follows:

$$\left.\begin{aligned} s_1 &= e^{-\alpha_1} \\ s_2 &= \exp\left[-\alpha_2 + j\left(\frac{-2\pi}{3} + \delta_2\right)\right] \\ s_3 &= \exp\left[-\alpha_3 + j\left(\frac{2\pi}{3} + \delta_3\right)\right] \end{aligned}\right\} \alpha_i \geqq 0 \quad (i = 1, 2, 3). \quad (24)$$

In this description of the imperfect circulator, no element has been included that could alter the symmetry of the junction and therefore all the properties enumerated earlier in points 1)–4) still apply. In particular, the property expressed in point 4) will remain valid. By utilizing the relationships (16)–(18) when the eigenvalues are given by (24), one obtains for the elements of $\mathbf{S}^{(j)}$ the following relationships:

$$\begin{aligned} S' &= \frac{1}{3}\left\{e^{-\alpha_1} - \frac{1}{2}\left(e^{-\alpha_3 + j\delta_3} + e^{-\alpha_2 + j\delta_2}\right) + j\frac{\sqrt{3}}{2}\left(e^{-\alpha_3 + j\delta_3} - e^{-\alpha_2 + j\delta_2}\right)\right\} \\ S'' &= \frac{1}{3}\left\{e^{-\alpha_1} + e^{-\alpha_2 + j\delta_2} + e^{-\alpha_3 + j\delta_3}\right\} \\ S''' &= \frac{1}{3}\left\{e^{-\alpha_1} - \frac{1}{2}\left(e^{-\alpha_3 + j\delta_3} + e^{-\alpha_2 + j\delta_2}\right) - j\frac{\sqrt{3}}{2}\left(e^{-\alpha_3 + j\delta_3} - e^{-\alpha_2 + j\delta_2}\right)\right\}. \end{aligned} \quad (25)$$

If the α_i and the δ_i are small, we obtain, by expanding the exponentials in series,

$$\begin{aligned} S' &\cong \frac{1}{3}\left\{\left[-\alpha_1 + \frac{1}{2}(\alpha_3 + \alpha_2) - \frac{\sqrt{3}}{2}(\delta_3 - \delta_2)\right] - j\frac{1}{2}\left[\delta_3 + \delta_2 + \sqrt{3}(\alpha_3 - \alpha_2)\right]\right\} \\ S'' &= \frac{1}{3}\left\{\left[3 - (\alpha_1 + \alpha_2 + \alpha_3)\right] + j[\delta_2 + \delta_3]\right\} \\ S''' &\cong \frac{1}{3}\left\{\left[-\alpha_1 + \frac{1}{2}(\alpha_3 + \alpha_2) + \frac{\sqrt{3}}{2}(\delta_3 - \delta_2)\right] - j\frac{1}{2}\left[\delta_3 + \delta_2 - \sqrt{3}(\alpha_3 - \alpha_2)\right]\right\}. \end{aligned} \quad (26)$$

By the inspection of (25) and (26), and remembering that $|S'|^2$, $|S''|^2$, and $|S'''|^2$ correspond respectively to the power reflected at the input arm, the power transmitted at the forward arm, the power transmitted at the isolated arm, one derives the following observations: 1) for a lossless circulator, affected by small values of δ_2 and δ_3, $|\delta'|^2 = |S'''|^2$, *i.e.*, the power appearing at the isolated arm equals the power reflected at the input arm; 2) for a lossy circulator, in the general case $|S'|^2 \neq |S'''|^2$ even for small values of the α_i and δ_i ($i = 1, 2, 3$) and for suitable combinations of these last quantities, the ratio $|S'|^2/|S'''|^2$ can be quite different from unity.

Acknowledgment

The authors are greatly indebted to Dr. M. H. Sirvetz for his helpful advice and his continued interest in this work. Appreciation is also expressed for the technical assistance of J. Matsinger during the execution of the experimental measurements.

A Wide-Band UHF Traveling-Wave Variable Reactance Amplifier*

R. C. HONEY† AND E. M. T. JONES†

Summary—**The techniques developed for designing periodically loaded traveling-wave parametric amplifiers using variable-reactance diodes are described in detail. An amplifier was built and tested with two different sets of eight diodes. The performance of the amplifier with each set of diodes agrees substantially with the theoretical predictions, the measured noise figures being about 1.2 db higher than the theoretical values in each case. The gain of the second amplifier varied from a minimum of 6.7 db to more than 13 db over the band from 550 to 930 mc, with a measured noise figure of 2.3 db for wide-band noise inputs in the middle of the band, corresponding to about 4.9 db for single-frequency inputs.**

* Manuscript received by the PGMTT, December 7, 1959; revised manuscript received, January 28, 1960.

† Stanford Research Institute, Menlo Park, Calif.

General

TRAVELING-WAVE parametric amplifiers have a number of useful properties, such as wide bandwidth and unilateral amplification, that have been thoroughly discussed in the rapidly expanding litera-

ture on this subject.[1–21] The detailed design procedure and the preliminary experimental results on one such amplifier are reported here; a schematic diagram of the amplifier is shown in Fig. 1. This amplifier uses eight variable-reactance diodes periodically loading a uniform high-impedance TEM transmission line. As shown in Fig. 1, this line propagates both the signal and the idling frequencies. An inductive reactance consisting of a suitable length of short-circuited transmission line is connected across each diode, forming a parallel resonant circuit whose resonant frequency is adjusted to fall within the first pass band of the periodically loaded line.

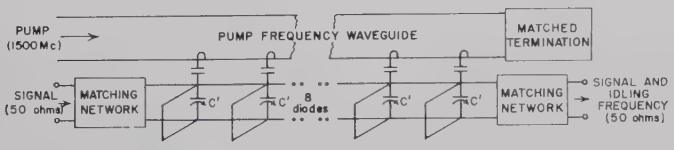


Fig. 1—Traveling-wave parametric amplifier.

[1] R. Adler, "A New Principle of Signal Amplification," presented at the Conf. on Electron Tube Research, Berkeley, Calif.; June, 1957.
[2] Y. Miyakawa, "Amplification and frequency conversion in propagating circuits," *Inst. Elec. Comm. Engrs. Japan Natl. Conv. Record*, p. 8; November, 1957. (In Japanese.)
[3] A. L. Cullen, "A traveling-wave parametric amplifier," *Nature*, vol. 181, p. 332; February 1, 1958.
[4] P. K. Tien and H. Suhl, "A traveling-wave ferromagnetic amplifier," PROC. IRE, vol. 46, pp. 700–706; April, 1958.
[5] W. H. Louisell and C. F. Quate, "Parametric amplification of space charge waves," PROC. IRE, vol. 46, pp. 707–716; April, 1958.
[6] A. Uhlir, Jr., "The potential of semiconductor diodes in high-frequency communications," PROC. IRE, vol. 46, p. 1099; June, 1958.
[7] R. Adler, "Parametric amplification of the fast electron wave," PROC. IRE, vol. 46, pp. 1300–1301; June, 1958.
[8] L. D. Buchmiller and G. Wade, "Pumping to extend traveling-wave tube frequency range," PROC. IRE, vol. 46, pp. 1420–1421; July, 1958.
[9] P. K. Tien, "Parametric amplification and frequency mixing in propagating circuits," *J. Appl. Phys.*, vol. 29, pp. 1347–1357; September, 1958.
[10] R. S. Engelbrecht, "A low-noise nonlinear reactance traveling-wave amplifier," PROC. IRE, vol. 46, p. 1655; September, 1958.
[11] R. Adler, G. Hrbek, and G. Wade, "A low-noise electron-beam parametric amplifier," PROC. IRE, vol. 46, pp. 1756–1757; October, 1958.
[12] A. Ashkin, "Parametric amplification of space charge waves," *J. Appl. Phys.*, vol. 29, pp. 1646–1651; December, 1958.
[13] G. Wade and R. Adler, "A new method for pumping a fast space-charge wave," PROC. IRE, vol. 47, pp. 79–80; January, 1959.
[14] W. E. Danielson, "Low noise in solid state parametric amplifiers at microwave frequencies," *J. Appl. Phys.*, vol. 30, pp. 8–15; January, 1959.
[15] R. S. Engelbrecht, "Nonlinear-Reactance Traveling-Wave Amplifiers for UHF," presented at Solid-State Circuits Conf., Philadelphia, Pa.; February, 1959.
[16] L. S. Nergaard, "Nonlinear-capacitance amplifiers," *RCA Rev.*, vol. 20, pp. 3–17; March, 1959.
[17] P. P. Lombardo and E. W. Sard, "Low-frequency prototype traveling-wave reactance amplifier," PROC. IRE, vol. 47, pp. 995–996; May, 1959.
[18] G. M. Roe and M. R. Boyd, "Parametric energy conversion in distributed systems," PROC. IRE, vol. 47, pp. 1213–1218; July, 1959.
[19] K. Kurokawa and J. Hamaski, "Mode theory of lossless periodically distributed parametric amplifiers," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-7, pp. 360–365; July, 1959.
[20] P. P. Lombardo and E. W. Sard, "Low-noise microwave reactance amplifiers with large gain-bandwidth products," 1959 IRE WESCON CONVENTION RECORD, pt. 1, pp. 83–98.
[21] C. V. Bell and G. Wade, "Circuit considerations in traveling-wave parametric amplifiers," 1959 IRE WESCON CONVENTION RECORD, pt. 2, pp. 75–82.

The pump frequency at 1500 mc is carried by a waveguide paralleling the TEM transmission line, and coupled loosely to each diode with loop coupling from the waveguide and capacitive coupling to the TEM line. The characteristics of the periodically loaded TEM line are chosen so that the pump frequency is in a stopband of the line. The upper sideband, at the signal frequency plus the pump frequency, is either in a stopband or is grossly out of phase for all signal frequencies, so that it cannot affect the interaction appreciably. Similarly, the signal and idling frequencies cannot propagate along the pump channel since they are below the cutoff frequency of the pump waveguide. The matching networks shown in the figure transform the high image impedance of the amplifier to an approximate 50-ohm level for use with conventional coaxial instrumentation.

Analysis and Design

General

In a periodically loaded transmission structure such as this, the conditions that must be satisfied to permit parametric interaction between the signal frequency and the pump frequency are given by

$$\omega_1 + \omega_2 = \omega_3, \tag{1}$$

$$\beta_1 + \beta_2 \approx \beta_3 \pm 2n\pi, \tag{2}$$

where

$n = 0, 1, 2, \cdots$
$\omega_1 = 2\pi f_1$, radians per second
f_1 = signal frequency, cycles per second
$\omega_2 = 2\pi f_2$, radians per second
f_2 = idling frequency, cycles per second
$\omega_3 = 2\omega f_3$, radians per second
f_3 = pump frequency, cycles per second
$\beta_m = \omega_m / v_m (m = 1, 2, 3)$
v_m = phase velocity of mth frequency along the loaded transmission line.

The idling frequency, ω_2, is generated by the variable capacitors so as to satisfy (1) exactly. However, (2) need be satisfied only approximately, since the gain of the amplifier is not greatly reduced if the phase error is small.[4]

The maximum gain for such an amplifier is given approximately by[21]

$$\alpha = (\omega_1\omega_2)^{1/2}(Z_{\pi 1}Z_{\pi 2})^{1/2}\Delta C/4 \text{ nepers per section}$$

$$= 13.65 Z_0 (f_1 f_2)^{1/2} \left(\frac{Z_{\pi 1}}{Z_0}\cdot\frac{Z_{\pi 2}}{Z_0}\right)^{1/2} \Delta C \text{ db per section} \tag{3}$$

where

ΔC = one-half the peak-to-peak capacitance variation ($C' = C + \Delta C \cos \omega_3 t$) in farads
C = average capacitance in farads
C' = total instantaneous capacitance in farads

$Z_{\pi 1}$ = image impedance in ohms of loaded line as seen by each diode at ω_1

$Z_{\pi 2}$ = image impedance in ohms of loaded line as seen by each diode at ω_2

Z_0 = characteristic impedance in ohms of uniform line.

This equation contains the most important factors which must be considered in the design of a traveling-wave parametric amplifier.

The following subsections will discuss in detail the analysis and design of circuits satisfying (1) and (2), methods for maximizing the gain in (3), and the design of the matching sections at each end of the amplifier.

Image Parameters of Periodically-Loaded Lines

When any transmission structure is shunt-loaded with variable-reactance diodes, the gain of the amplifier becomes proportional to the square root of the product of the image impedances of the structure at the signal and idling frequencies as seen at the terminals across which the diodes are connected, as given by (3). It turns out that the image impedance of a uniform transmission line with periodic shunt reactive elements of any type assumes a particularly simple form, and a knowledge of the phase velocity along the structure as a function of frequency, such as contained in the conventional ω vs β curve for that structure, is sufficient to predict immediately the image impedance of the structure. This can be shown in the following way. Assume that a uniform transmission system of characteristic admittance, Y_0, is periodically loaded with shunt susceptances, $j2B$, as shown in Fig. 2(a). Let the phase constant of the unloaded line be β_0, or $\beta_0 = 2\pi/\lambda = \omega/v$, where ω is the angular frequency, v the velocity of light in the medium, and λ the wavelength on the uniform transmission line. A typical half section of this line is shown in Fig. 2(b), for which the image phase constant, ϕ, in radians per section, can be found from either

$$\tanh \frac{\phi}{2} = \sqrt{\frac{Y_{oc}}{Y_{sc}}}, \tag{4}$$

or

$$\cosh \phi = \frac{Y_{sc} + Y_{oc}}{Y_{sc} - Y_{oc}}, \tag{5}$$

where

Y_{sc} = input admittance when opposite terminal-pair is short circuited

Y_{oc} = input admittance when opposite terminal-pair is open circuited.

The image admittances, Y_π and Y_T, are also found from the usual formula

$$Y_{\pi,T} = \sqrt{Y_{oc} Y_{sc}}. \tag{6}$$

The open- and short-circuit admittances for the left-hand terminal pair of the half section shown in Fig. 2(b) are given by

$$Y_{oc} = jB + jY_0 \tan (\beta_0 L/2)$$
$$Y_{sc} = jB - jY_0 \cot (\beta_0 L/2). \tag{7}$$

Hence, from (5),

$$\frac{B}{Y_0} = \frac{\cos \beta_0 L - \cos \beta L}{\sin \beta_0 L}, \tag{8}$$

where $\phi = j\beta L$ in the pass bands of the structure.

Finally, the π-section image admittance at the left-hand terminal pair Y_π, found by combining (6) to (8), reduces to the simple equation

$$\frac{Y_\pi}{Y_0} = \frac{\sin \beta L}{\sin \beta_0 L} \quad \text{or} \quad \frac{Z_\pi}{Z_0} = \frac{\sin \beta_0 L}{\sin \beta L}. \tag{9}$$

Eq. (9) can be plotted on an ω vs β diagram as shown in Fig. 3, forming an auxiliary grid of lines which are labeled with the constant values of Z_π/Z_0 along each line. Hence, if the ω vs β curve of any periodically shunt-

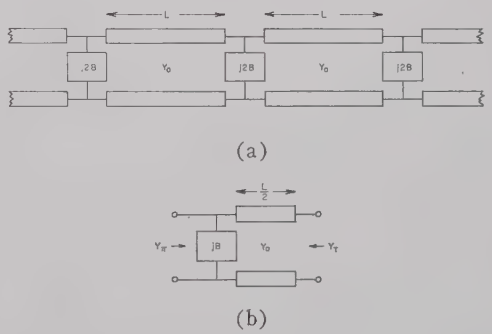


Fig. 2—Notation for periodically loaded transmission lines. (a) Periodically loaded transmission line. (b) Typical half section.

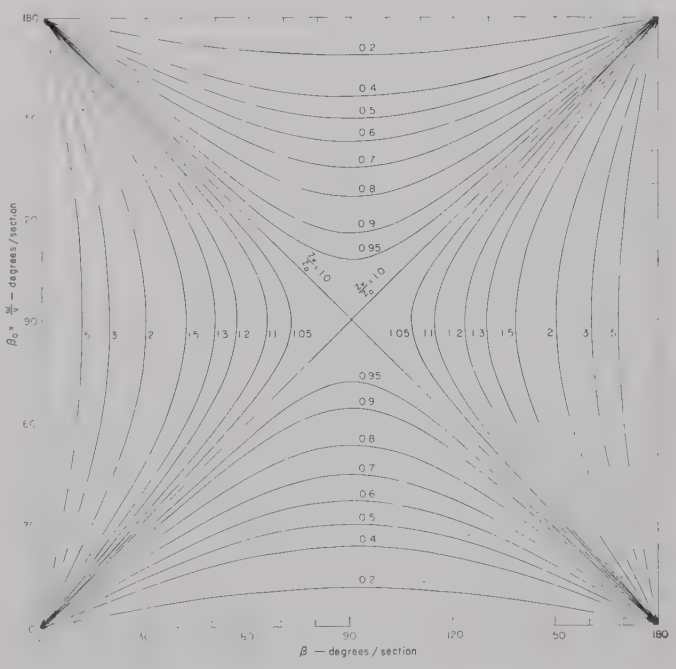


Fig. 3—Image impedances for periodic shunt-loaded transmission line.

loaded nondispersive transmission line is superimposed on this grid, the image impedances can be read directly from the intersections of the curve with the grid lines. This is true regardless of the nature of the reactive shunt loading. Since the gain of a parametric amplifier of the type considered here is proportional to the square root of the product of the image impedances at the signal and idling frequencies, it becomes a simple matter to compare the gains of two circuits of this type from their respective ω vs β curves, other factors being constant.

A study of Fig. 3 shows that maximum image impedances are obtained in circuits whose ω vs β curves pass through the center of the figure at $\beta_0 = \omega/v = \beta = 90$ degrees per section.

Periodic Structure Used in Amplifier

The actual structure chosen for use in the traveling-wave parametric amplifier described here is shown in Fig. 1, and uses a length of shorted transmission line of characteristic impedance Z_0 in parallel with the lumped capacitors to periodically load a uniform TEM transmission line. The shorted length of line resonates the average value of the shunt capacitor at 815 mc.

The TEM transmission line chosen for this amplifier consists of a small round wire over a ground plane. The ground plane allows convenient access to one end of the crystal diodes to build an RF bypass network so that dc bias can be applied to each diode individually. The spacing of the wire over the ground plane, $h = 0.308$ inch, was determined by the length of the ceramic insulation in the diode cartridge. The characteristic impedance was then determined by the choice of wire size, $d = 0.010$ inch, the smallest practical size that could be easily handled. For the geometry shown in Fig. 4 this gives[22]

$$Z_0 = 285.4 \text{ ohms},$$

which will be the value used in all subsequent calculations.

The ω vs β curve is plotted in Fig. 5 along with the position of the pump signal at 1500 mc, and the other important frequencies associated with the amplifier. The straight line with 45-degree slope represents the ω vs β characteristic of the unperturbed TEM line.

In this particular structure, the ω vs β curve passes through the degenerate frequency point, 750 mc, as shown in Fig. 5, permitting the maximum possible gain for 750 mc since (2) is satisfied exactly. At other frequencies, the phase errors that are introduced reduce the gain as the edges of the band are approached. This tends to compensate for the fact that the image impedances increase towards the edges of the band, causing the gain to increase.

It should be pointed out that the pump frequency should not be phased so that it lies on the lines $\beta = n\pi$, $n = 0, 1, 2, \cdots$. If it did, the amplifier would have the same gain in both the forward and the backward directions, eliminating one of the principal advantages of the traveling-wave parametric amplifier, namely, unilateral amplification. The placement of the pump frequency as shown in Fig. 5 ensures that an adequate phase error exists between the pump wave and the backward waves on the amplifier so that a relatively small amount of interaction will occur.

Note also from Fig. 5 that the upper sideband frequencies are either in a stopband of the periodically loaded line or are grossly misphased with respect to the pump signal, so that no appreciable interaction will take place.

The maximum gain that can be obtained from this structure is given by (3). This maximum total gain for an eight-section amplifier (assuming that there are no phase errors and that the structure is terminated in its image impedance) is shown in Fig. 6 for several values of the peak capacitance variation, ΔC.

The RF bypass networks, which allow individual dc bias to be applied to each diode separately, consist of a

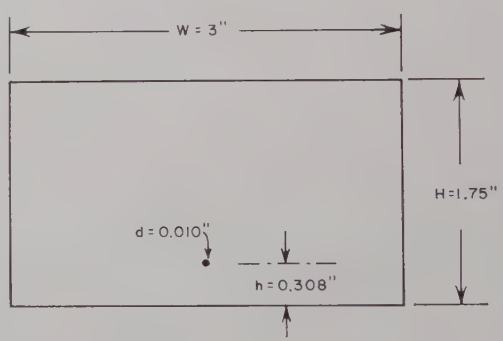


Fig. 4—Actual geometry of uniform transmission line.

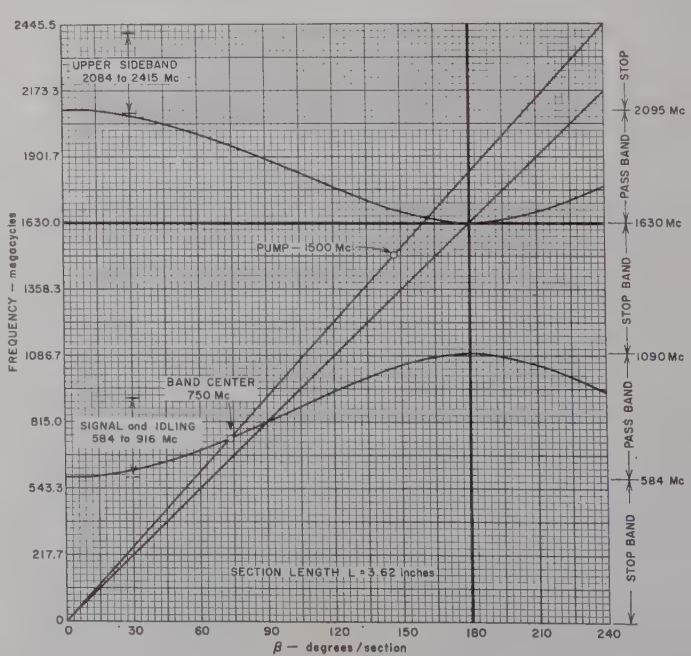


Fig. 5—ω vs β diagram for complete amplifier.

[22] "Reference Data for Radio Engineers," Federal Telegraph and Telephone Corp., New York N. Y., 4th ed., p. 593; 1956.

single π-section low-pass filter. The shunt capacitors are simply radial-line capacitors with thin mica dielectric, and the series inductance is a half-inch length of very high impedance coaxial line. No RF leakage could be detected past these filters, and no appreciable losses in the RF circuitry could be attributed to them.

Matching Networks

The purpose of the matching transformers shown in Fig. 1 is simply to transform the high image impedance of the amplifier down to a 50-ohm level for use with conventional coaxial instrumentation, and to terminate each end of the amplifier with impedances that are as close as possible to the image impedance of the amplifier. The π-section image impedance is much more constant over the signal band than the T-section image impedance. Therefore, an extra section was added to each end of the amplifier as shown in Fig. 7 so that the matching transformers could look into this π-section impedance. An average value of the π-section image impedance was then selected (336 ohms in this case) and an optimally designed stepped-transmission-line transformer[23] was used to match from 50 ohms into 336 ohms. This design used two intermediate quarter-wave sections ($n=3$) for a design bandwidth of 1.4 to 1. The end sections and matching transformers are shown with their important dimensions in Fig. 7.

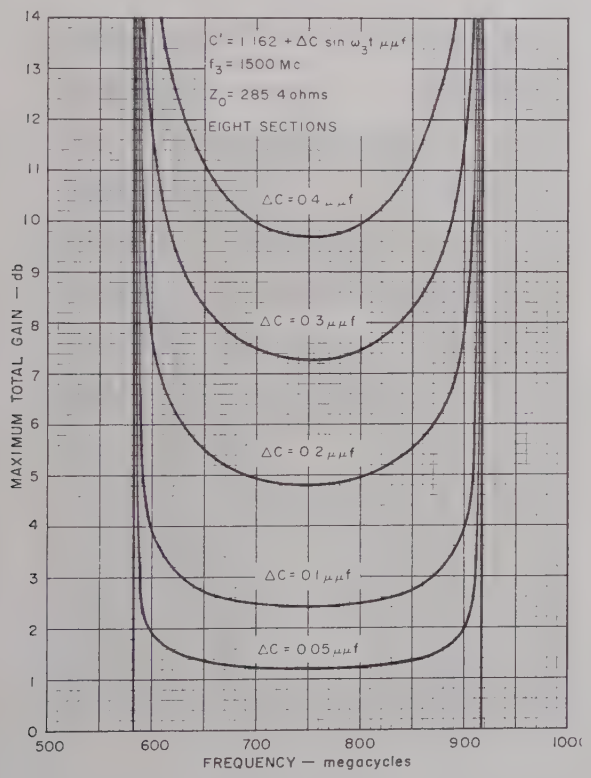


Fig. 6—Theoretical maximum total gain vs frequency.

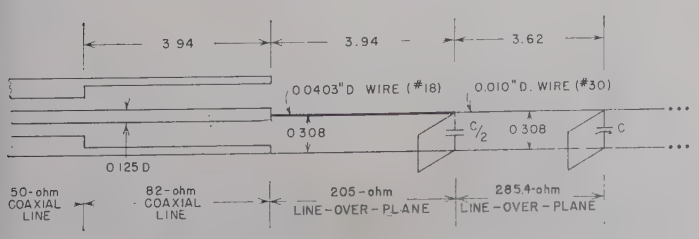


Fig. 7—Matching networks.

The resultant over-all measured VSWR looking into the amplifier when it is terminated in a matched 50-ohm load is below 1.5:1 from 650 to 900 mc, and below 2:1 from 610 to 960 mc. The VSWR rises rapidly outside this band as the cutoff frequencies are approached.

In practice it was found that optimum performance was obtained if the networks were slightly modified by removing the parallel inductances adjacent to the amplifier terminals and the fixed capacitors at these terminals were replaced with variable capacitors that could be adjusted externally. These changes slightly improved the matches looking into the amplifier, keeping the measured input VSWR below 2:1 from 620 to 1020 mc.

Test Cavity

In general, the available crystal diodes are not identical, and significant differences occur from diode to diode. Therefore, it was decided to test each diode individually in a separate cavity before installing it in the traveling-wave amplifier. This test cavity consisted of a length of transmission line with the same geometry as that in the parametric amplifier (Fig. 4). The length of the cavity was chosen to be one-half wavelength long at the resonant frequency of the shunt circuit, 815 mc. At the center of this cavity, the diodes could be mounted in shunt across the transmission line in exactly the same manner as in the amplifier. The cavity was used as a transmission cavity, with very loose loop coupling at each end of the cavity to keep the insertion loss greater than 40 db for all measurements. With no diode or diode mount in place, the cavity resonates at 815 mc. With a diode in place, the resonant frequency was lowered due to the shunt capacitance of the diode and its mount. By measuring this frequency, the total effective capacitance can be very accurately determined as a function of the external bias, and by measuring the Q of the resonance, the losses in the diode and the RF bypass filter can be inferred.

The reduction in the resonant frequency due to a lumped capacitor at the center of the cavity can be simply calculated in the following way. At resonance, the sum of the admittances looking in either direction from the center equals zero; therefore

$$j\frac{B}{2} - jY_0 \cot\frac{\beta_0 \mathfrak{L}}{2} = 0, \tag{10}$$

[23] S. B. Cohn, "Optimum design of stepped transmission-line transformers," IRE Trans. on Microwave Theory and Techniques, vol. 3, pp. 16–21; April, 1955.

where, for this case,

$B=\omega C$, the susceptance of the lumped capacitance
$Y_0=1/Z_0$
$Z_0=285.4$ ohms, the characteristic impedance of the transmission line
$\beta_0=\omega/v=$ the phase constant of the unloaded transmission line
$\mathcal{L}=7.23$ inches, the length of the cavity ($\lambda/2$ at 815 mc).

Hence, the shunt capacitance can be found as a function of frequency from

$$C = \frac{2Y_0}{\omega}\cot\frac{\beta_0\mathcal{L}}{2} = \frac{2Y_0}{\omega}\cot\frac{\omega\mathcal{L}}{2v}. \tag{11}$$

The Q's of the diodes, Q_d, can also be found from the Q's of the cavity plus diode, Q_0, from

$$Q_d = 2Q_0\left(1+\frac{\beta_0\mathcal{L}}{\sin\beta_0\mathcal{L}}\right)^{-1} \tag{12}$$

assuming that the losses in the cavity are negligible compared to the losses in the diodes.[24]

After measuring the effective capacities and Q's of each diode as a function of bias, a length of transmission line (perpendicular to the main transmission line) was connected in parallel with the diode. This added line has a short circuit that can be adjusted to couple the desired inductive reactance across the diode. The position of this short circuit was adjusted until the cavity again resonated at 815 mc, at which frequency the parallel combination of short-circuited line plus diode and mount were also in resonance, resulting in a high resistive load across the transmission line in the cavity at the resonant frequency. The lengths of short-circuited line required to resonate each diode at a given bias were carefully measured, and the short circuits in the traveling-wave amplifier were adjusted to the same values.

Theoretical Noise Figure

Without Dissipation Loss: When a signal of frequency f_1 is applied to the input of a traveling-wave variable-reactance amplifier, the signal increases in amplitude as it travels down the length of the amplifier. At each diode the signal frequency, f_1, mixes with the pump frequency, f_3, and generates two principal sidebands at f_3+f_1 and f_3-f_1. The power in the lower sideband, or idling frequency, at $f_2=f_3-f_1$ also increases as it travels down the amplifier, while the power in the upper sideband interferes destructively at each diode.

If a current of amplitude a is applied to the left-hand, or input, end of the amplifier, one finds that the current on the transmission line to the right of the nth diode is given approximately by

$$I_1(n,t) = a\cosh(n\alpha)e^{j[2\pi f_1 t-(n-1)\beta_1]}. \tag{13}$$

Similarly, the current at the idling frequency f_2 just to the right of the nth diode is[25]

$$I_1^*(n,t) = \sqrt{\frac{f_2}{f_1}\frac{Z_{\pi 1}}{Z_{\pi 2}}}\,a\sinh(n\alpha)e^{j[2\pi f_2 t-(n-1)\beta_2]}. \tag{14}$$

In (14) the symbol * means complex conjugate. If one applies a current of amplitude a and frequency f_2 to the input, the currents to the right of each diode at frequencies f_1 and f_2 are determined by interchanging the subscripts in (13) and (14).

If a matched noise generator, having available noise power kT_1B at $T_1°$ K whose noise spectrum is centered about f_1, is connected to the input of the amplifier, it is clear that

$$kT_1B = I_1I_1^*Z_{\pi 1} = a^2Z_{\pi 1} \tag{15}$$

where Boltzmann's constant $k=1.38\times10^{-23}$ watt-seconds per degree Kelvin. Hence the noise power $N_1'(n)$ in the bandwidth B centered at f_1 to the right of the nth diode becomes

$$N_1'(n) = kT_1B\cosh^2(n\alpha), \tag{16}$$

while the noise power, $N_2'(n)$ in the bandwidth B centered at f_2 to the right of the nth diode becomes

$$N_2'(n) = kT_1B\frac{f_2}{f_1}\sinh^2(n\alpha). \tag{17}$$

Similar expressions are obtained when a matched generator at temperature T_2 having available noise power kT_2B, whose noise spectrum is centered about f_2, is connected to the input of the amplifier. Then the noise power, $N_1''(n)$, in the bandwidth, B, centered about f_1 to the right of the nth diode is

$$N_1''(n) = kT_2B\frac{f_1}{f_2}\sinh^2(n\alpha), \tag{18}$$

while the noise power, $N_2''(n)$, in the bandwidth, B, centered at f_2 becomes

$$N_2''(n) = kT_2\cosh^2(n\alpha). \tag{19}$$

Therefore, the total noise powers out of the signal and idling frequency channels are given, respectively, by

$$N_1(n) = N_1'(n) + N_1''(n) = kB\left[T_1\cosh^2(n\alpha) + T_2\frac{f_1}{f_2}\sinh^2(n\alpha)\right] \tag{20}$$

[24] E. M. T. Jones, G. L. Matthaei, S. B. Cohn, and B. M. Schiffman, "Design Criteria for Microwave Filters and Coupling Structures," Stanford Res. Inst., Menlo Park, Calif., Tech. Rept. 5, Contract DA 36-039 SC-74862, pp. 5–11; March, 1959.

[25] Eqs. (13) and (14) are obtained in an obvious way from (27) and (28) of Tien and Suhl, *op. cit.*

$$N_2(n) = N_2'(n) + N_2''(n)$$

$$= kB\left[T_2 \cosh^2(n\alpha) + T_1 \frac{f_2}{f_1} \sinh^2(n\alpha)\right]. \quad (21)$$

Eqs. (20) and (21) can also be applied to signals of power P_1 in Channel 1 and P_2 in Channel 2, in which case the output signal powers in the two channels are given by

$$S_1(n) = P_1 \cosh^2(n\alpha) + P_2 \frac{f_1}{f_2} \sinh^2(n\alpha) \quad (22)$$

$$S_2(n) = P_2 \cosh^2(n\alpha) + P_1 \frac{f_2}{f_1} \sinh^2(n\alpha). \quad (23)$$

The noise figure of an amplifier is defined by

$$F = \frac{N_{\text{out}}}{N_{\text{in}}} \cdot \frac{S_{\text{in}}}{S_{\text{out}}} \quad (24)$$

where N_{in} is the available input thermal-noise power from the generator resistance (at a temperature $T_1 = 290°\text{K}$ by convention) in the band around the signal frequency, N_{out} is the available output noise power, S_{in} is the available input signal power from the generator, and S_{out} is the available output signal power. $S_{\text{out}}/S_{\text{in}}$ may be defined as the power gain of the amplifier. Because of the multiplicity of channels that must be considered in amplifiers such as this, several different modes of operation can be visualized which may have different *effective* noise figures.

These different modes can be treated as follows, using (20) to (24):

Case I—Input signal on Channel 1 only; output detected on Channel 1 only.

$$(P_2 = 0,\ T_1 = 290°\text{K})$$

$$F_{\text{I}} = \frac{N_1(n)}{k290B} \cdot \frac{P_1}{S_1(n)}$$

$$= 1 + \frac{T_2}{290} \frac{f_1}{f_2} \tanh^2(n\alpha)$$

$$\approx 1 + \frac{T_2}{290} \frac{f_1}{f_2} \text{ for large gain } (n\alpha > 2). \quad (25)$$

This case corresponds to the most usual manner for using these amplifiers with single-frequency signals, and the noise figure so found corresponds to the IRE definition of noise figure.[26] The manner in which this noise figure varies with gain is illustrated in Fig. 8.

Case II—Input signal on Channel 1 only; output detected on Channel 2 only.

$$(P_2 = 0,\ T_1 = 290°\text{K})$$

[26] "Standards on Receivers: Definitions of Terms, 1952," PROC. IRE, vol. 40, pp. 1681–1685; December, 1952.

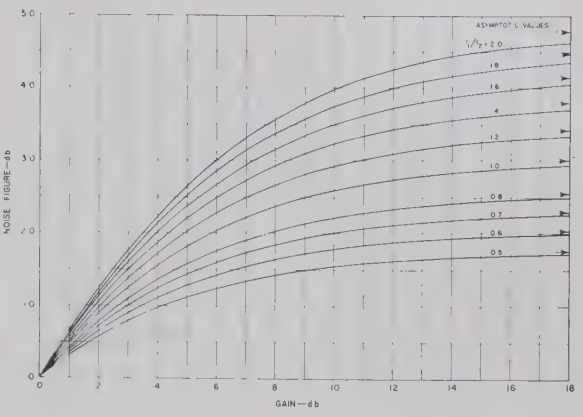


Fig. 8—Noise figure vs signal gain.

$$F_{\text{II}} = \frac{N_2(n)}{k290B} \cdot \frac{P_1}{S_2(n)}$$

$$= 1 + \frac{T_2}{290} \frac{f_1}{f_2} \coth^2(n\alpha)$$

$$\approx 1 + \frac{T_2}{290} \frac{f_1}{f_2} \text{ for large gain } (n\alpha > 2). \quad (26)$$

This case is always worse than Case I, although they approach each other for high gain. Again, the IRE definition of noise figure agrees with this definition.

Case III—Input signal on Channel 1 only; output detected on Channels 1 and 2.

$$(P_2 = 0,\ T_1 = 290°\text{K})$$

$$F_{\text{III}} = \frac{N_1(n) + N_2(n)}{k290B} \cdot \frac{P_1}{S_1(n) + S_2(n)}$$

$$= 1 + \frac{T_1}{290} \frac{1 + \frac{f_1}{f_2} \tanh^2(n\alpha)}{1 + \frac{f_2}{f_1} \tanh^2(n\alpha)}$$

$$\approx 1 + \frac{T_2}{290} \cdot \frac{f_1}{f_2} \text{ for large gain } (n\alpha > 2). \quad (27)$$

There are regions in which Case III will yield a lower noise figure than Case I, and others where it will be higher, although they approach each other for high gain.

Case IV—Uncorrelated input signals on Channels 1 and 2; output detected on Channel 1 only.

$$(P_1 = P_2 = P,\ T_1 = T_2 = 290°\text{K})$$

$$F_{\text{IV}} = \frac{N_1(n)}{2k290B} \cdot \frac{2P}{S_1(n)} = 1. \quad (28)$$

Case V—Uncorrelated input signals on Channels 1 and 2; output detected on Channels 1 and 2.

$$(P_1 = P_2 = P,\ T_1 = T_2 = 290°\text{K})$$

$$F_V = \frac{N_1(n) + N_2(n)}{2 \times k290B} \cdot \frac{2P}{S_1(n) + S_2(n)} = 1. \quad (29)$$

This case and Case IV correspond to the usual manner for using these amplifiers with broad-band noise signals, such as in many radio astronomy applications and in measuring the noise figures of these amplifiers with broad-band gas-discharge, noise sources. Case V also corresponds to the use of these amplifiers around the degenerate frequency by considering that the total bandwith B' of the amplifier is divided into a bandwidth B just below the degenerate frequency, $f_2/2$, and an equal bandwidth B just above the degenerate frequency.

Case VI—Correlated input signals on Channels 1 and 2; output detected on Channels 1 and 2; degenerate case.

($f_1 = f_2$ and phased so that signal voltages add)

$$(P_1 = P_2 = P,\ T_1 = T_2 = 290°\text{K})$$

$$F_{VI} = \frac{N_1(n) + N_2(n)}{2 \times k290B} \cdot \frac{2P}{[\sqrt{S_1(n)} + \sqrt{S_2(n)}]^2} = \frac{1}{2}. \quad (30)$$

This rather surprising answer[27] results from the fact that the phase of the signal has been assumed to be locked to the phase of the pump in such a way that maximum gain occurs, *i.e.*, that more information (phase information) is required to make this system work.

With Dissipation Loss: In approximating the noise figure of an n-section traveling-wave variable-reactance amplifier whose diodes have dissipation loss, it is convenient to consider the amplifier as a cascade of n amplifiers with noise figures, F_n, and cold insertion losses, L_n. If the input and output bandwidths include both signal and idle channels (Case V), the over-all noise figure of the complete amplifier, F_V, can be found from the total output noise power

$$N_{out} = k290B\left[\cosh^2(n\alpha) + \frac{1}{2}\left(\frac{f_1}{f_2} + \frac{f_2}{f_1}\right)\sinh^2(n\alpha)\right]F_V$$

$$= k290B\left\{F_1\left[\cosh^2(n\alpha) + \frac{1}{2}\left(\frac{f_1}{f_2} + \frac{f_2}{f_1}\right)\sinh^2(n\alpha)\right]\right.$$

$$+ (F_2 - 1)\left[\cosh^2(n-1)\alpha\right.$$

$$\left. + \frac{1}{2}\left(\frac{f_1}{f_2} + \frac{f_2}{f_1}\right)\sinh^2(n-1)\alpha\right] + \cdots$$

$$+ (F_n - 1)\left[\cosh^2\alpha\right.$$

$$\left.\left. + \frac{1}{2}\left(\frac{f_1}{f_2} + \frac{f_2}{f_1}\right)\sinh^2\alpha\right]\right\}, \quad (31)$$

where the α's are computed from the total net gain of the amplifier,

$$\cosh^2(n\alpha) + \frac{1}{2}\left(\frac{f_1}{f_2} + \frac{f_2}{f_1}\right)\sinh^2(n\alpha). \quad (32)$$

If one operates the amplifier in the vicinity of the degenerate frequency, so that $f_1 \approx f_2$, then (31) reduces to the simpler expression

$$F_V - 1 = (F_1 - 1) + (F_2 - 1)\frac{\cosh 2(n-1)\alpha}{\cosh 2n\alpha}$$

$$+ \cdots + (F_n - 1)\frac{\cosh 2\alpha}{\cosh 2n\alpha} \quad (33)$$

where the α's are computed from the total *net* gain of the amplifier,

$$\cosh^2 n\alpha + \sinh^2 n\alpha = \cosh(2n\alpha). \quad (34)$$

These expressions differ from the usual cascaded-amplifier noise figure expressions since the gain of any given stage for uncorrelated noise generated within itself is, in general, less than its gain for the correlated noise and signals received from the previous stage.

Assuming that the losses are the same at f_1 and f_2, the noise figure of each stage, F_n, is approximately equal to its insertion loss, L_n, for small gain per stage. Since the losses in each stage are also assumed to be equal,

$$L_1 = L_2 = \cdots = L_n = L^{1/n},$$

where L is the total insertion loss in the amplifier expressed as a power ratio ≥ 1. Therefore, the total excess noise from the amplifier operated near the degenerate frequency is given by

$$F_V - 1 = \frac{L^{1/n} - 1}{\cosh 2n\alpha}[\cosh 2n\alpha + \cosh 2(n-1)\alpha$$

$$+ \cdots + \cosh 2\alpha]. \quad (35)$$

Assuming that the noise in the output of the amplifier divides up in about the same way between the signal and idling channels in the lossy case as in the lossless case, the noise figure for single-channel operation (Case I) can be found from the double-channel operation case treated above by comparing (25) and (29) to give

$$F_I \approx F_V\left(1 + \frac{f_1}{f_2}\tanh^2 n\alpha\right) \quad (36)$$

where F_V can be found from (31).

Experimental Procedure and Results

Diodes in Test Cavity

A variety of diodes were checked in the test cavity described in the previous section. The eight diodes chosen for the tests of the traveling-wave parametric amplifier were selected from a group of diffused-junction silicon diodes from Microwave Associates, Inc., and Bell Telephone Laboratories. The total effective capaci-

[27] S. B. Cohn, "The noise figure muddle," *Microwave J.*, vol. 2, pp. 7–11; March, 1959.

ties of these diodes plus their mounts as a function of external dc bias are shown in Fig. 9.

An inductance consisting of a length of short-circuited transmission line was then connected in parallel with each diode and adjusted as described previously. The final over-all Q_0's of the test cavity with diodes and short-circuited lines varied from 90 to 120. Assuming that all of the losses occur in the series resistance of the diodes, the actual Q of the diodes, Q_d, can be found from (12), giving Q_d's around 50.

Amplifier

Operation: A photograph of the amplifier is shown in Fig. 10. The variable-width pump waveguide loaded with a $1\frac{1}{4}$-inch-diameter polystyrene rod placed on top of a $1\frac{1}{2}$-inch-high by 1-inch-wide metal ridge is seen on the left. This composite loading structure was used in order to achieve the proper phase velocity of the 1500-mc TE_{10} mode in the pump waveguide and at the same time to suppress in the pump waveguide the TE_{20} mode of propagation at the pump frequency of 1500 mc and to suppress the TE_{10} mode of propagation in the pump waveguide at 916 mc, the highest signal or idling frequency. The waveguide and the adjustable probes coupling to each diode are also evident in the photograph. The matching networks can be seen at either end of the amplifier. The RF bypass networks are contained in the structure along the right edge of the amplifier. External to these bypass networks, a 1000-micromicrofarad condenser and a 5-megohm variable resistance were connected to ground across each diode. The resistors provide dc return paths for the diodes, and a resistance across which self-bias voltage can be developed. The capacitors provide some measure of protection for the diodes from static discharges and other external transients that might damage them.

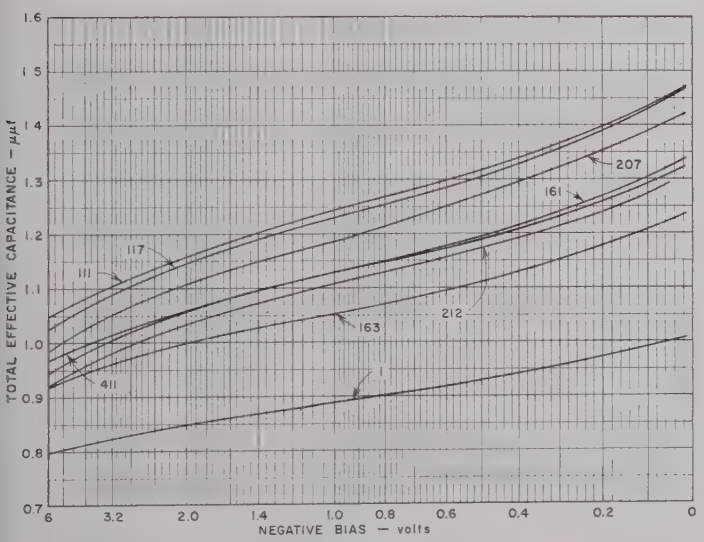


Fig. 9—Total capacitance vs bias.

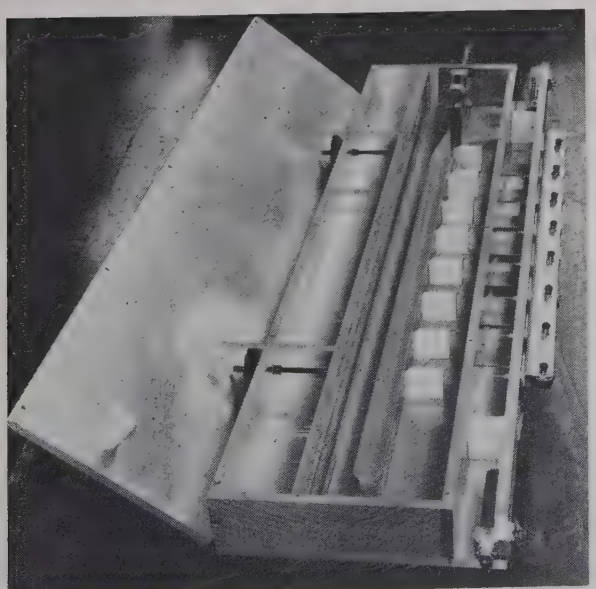


Fig. 10—Completed traveling-wave parametric amplifier.

The procedure for tuning up this amplifier was as follows. First, the couplings from the pump waveguide to each diode were adjusted so that the loop sizes and the gaps forming the capacitive couplings to each diode were all mechanically identical. Then, the individual bias resistors connected across each diode were adjusted to provide optimum performance without oscillation anywhere in the band.

The 1500-mc CW pump power was supplied by an Airborne Instruments Laboratory Model 124A Power Oscillator followed by suitable filters, providing at least 10 watts of pump power to the pump waveguide when required. Only a very small fraction of this total power was coupled out to the diodes, so that the actual pump power required by the amplifier could not be readily measured.

Normally, the amplifier was fed from a well-padded 1000-cycle-modulated signal generator at sufficiently low signal levels that the self-bias voltages were not affected by the signal even in regions of high gain. The amplifier was terminated with a large, well-matched attenuator followed by a low-pass filter that passed the signal and idling frequencies but was cut off to the pump frequency. Both the signal and the idling frequencies were detected simultaneously with a matched broadband detector feeding a narrow-band 1000-cycle amplifier.

The input signal generator was mechanically swept over the band from about 500 mc to 930 mc in synchronism with the horizontal sweep of an oscilloscope with a long-persistence screen. The output from the square-law detector was fed to the vertical deflection plates, and a continuous plot of total output power vs frequency was obtained on the oscilloscope. It was then a very simple matter to make adjustments to the amplifier and observe their effects everywhere in the band. When a good combination of the various parameters was found, a photograph of the trace on the oscilloscope recorded all the data. These photographs, combined with similar traces of the output of the detector when

the pump signal was removed, and also when the amplifier was bypassed, were sufficient to provide a fairly complete picture of the amplifier performance. Point-by-point measurements were also taken of a few of the better operating points to check the photographic data.

Gain: The final adjustments to the amplifier to obtain optimum response yielded the experimental results shown by the solid curves in Fig. 11. Because of the relatively high losses in the diodes, the cold insertion loss of the amplifier was about 3 db or more over most of the band. To measure the cold insertion loss, the external bias voltage applied to the diodes was set equal to the self-bias voltage on the diodes when they are being pumped. This high insertion loss limited the net signal gain that could be obtained from the amplifier to a minimum of about 4 db over most of the band. Attempts to increase this gain at the middle of the band only resulted in oscillation at the edges of the band where the gain is very high due to the high image impedances and regeneration within the amplifier.

The solid curves in Fig. 11 are the output of the broad-band detector, corresponding to the sum of the signal and idling frequency powers. Since, in general, these two powers are not equal, the gain in the signal channel alone can be calculated approximately in the following way. As shown previously, the power in the signal channel increases as $\cosh^2(n\alpha)$ along the amplifier, and in the idling channel as $(f_2/f_1)\sinh^2(n\alpha)$. The sum of these powers is then proportional to

$$P_T = \cosh^2(n\alpha) + \frac{f_2}{f_1}\sinh^2(n\alpha)$$

which is the quantity measured by the broad-band detector assuming that it is equally sensitive to the powers at the two frequencies. This quantity is always greater than the power in the signal channel alone by the ratio

$$R = \frac{P_T}{P_S} = \frac{\cosh^2(n\alpha) + \frac{f_2}{f_1}\sinh^2(n\alpha)}{\cosh^2(n\alpha)}$$

$$= 1 + \frac{f_2}{f_1}\tanh^2(n\alpha). \qquad (37)$$

Thus, knowing the frequency ratio, f_2/f_1, and the total gain, the signal gain can be determined. This calculation has been applied to the data in Fig. 11 and the results are shown by the dashed curves in the figure.

It is apparent, however, that in certain applications, over-all system performance could be significantly improved by detecting both the signal and the idling frequency powers.

Noise Figure: The noise figure of the amplifier was measured using wideband noise generated by a gas-discharge noise source and fed into both the signal and the idling channels on each side of the degenerate frequency, $f_3/2$. The output of the amplifier in the band around the degenerate frequency was fed to a crystal-mixer, through a 30-mc IF amplifier to a Hewlett-Packard noise-figure meter that automatically computes the noise figure from the ratio of the signal it receives when the noise tube is on to the signal it receives when the noise tube is off.

The complete block diagram of the experimental equipment used to measure this noise figure near the degenerate frequency (750 mc) is shown in Fig. 12. The noise figure of the amplifier alone is computed in the usual way,

$$F_1 = F - \frac{F_2 - 1}{G_1} \qquad (38)$$

where F_1 and G_1 are the noise figure and the power gain of the parametric amplifier, F_2 is the measured noise figure of the second stage, and F is the measured over-all noise figure of both stages. These noise figures correspond to those already defined in Case V.

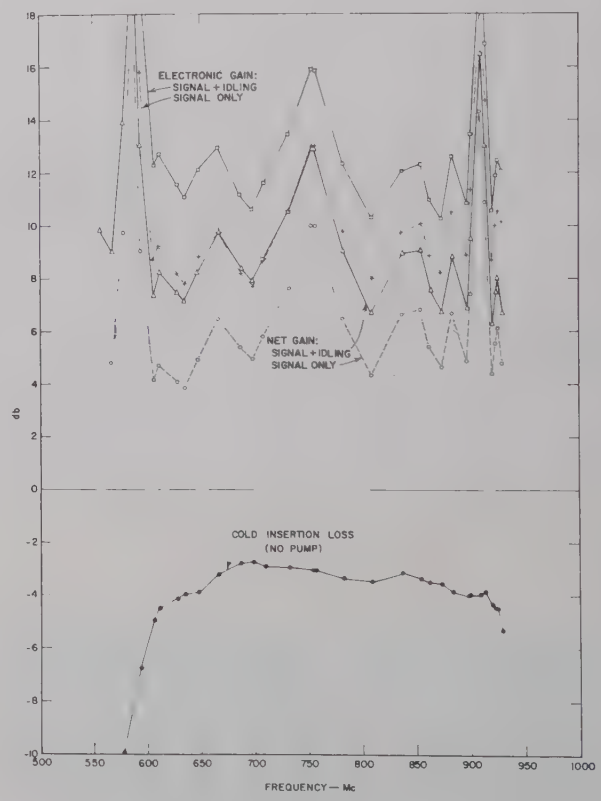


Fig. 11—Performance of amplifier.

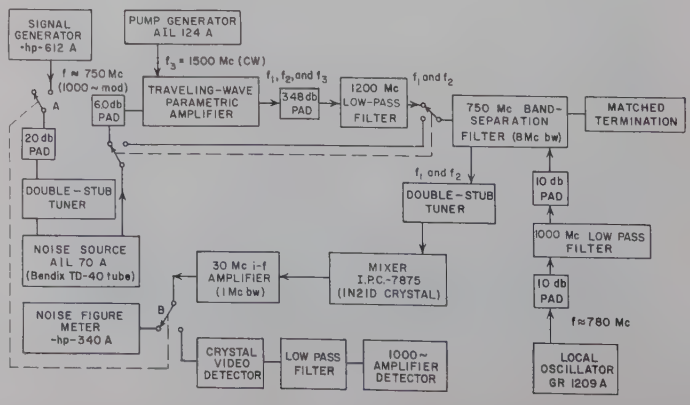


Fig. 12—Experimental setup for measuring noise figures.

The strip-line band-separation filter included in Fig. 12 was built to use in measuring the noise figure and is of the type shown in Fig. 3(e) in Cohn and Coale.[28] This type of filter has the advantage that it appears well-matched to the amplifier at all frequencies, thus greatly reducing the regeneration within the amplifier, yet it diverts an 8-mc band around the degenerate frequency from the output of the amplifier into the second stage. Furthermore, it provides a convenient means of introducing the local-oscillator signal into the crystal mixer as shown in the figure.

In using the gas-discharge noise source, and applying (38), a number of precautions must be taken to ensure accurate results. These precautions include such things as:

1) Corrections of the noise-figure readings due to the difference between the actual excess noise figure of the gas tube and its rated value of 15.2 db due to incomplete coupling (determined from its hot insertion loss at 750 mc, 22.4 db) and internal losses (determined from its cold insertion loss at 750 mc, 0.75 db).[29] For this particular tube at 750 mc this amounts to 0.35 db subtracted from all noise-figure meter readings.
2) Accurately measuring and accounting for all losses in the pads and cables used in the measuring set-up.
3) Ensuring that the output impedance of the noise source is exactly the same whether the tube is off or on. If this impedance were to change, it would change the phase or amplitude of any reflection from the noise source, hence changing the regeneration and the gain within the amplifier due to these small reflections. This is the purpose of the double stub tuner preceding the noise source shown in Fig. 12. The pads at the inputs to each amplifier also help to reduce this effect.
4) Ensuring that the temperature of the gas discharge tube does not rise sufficiently far above 290°K during operation to appreciably alter the noise from the source when the gas discharge is off. This noise is a function of both the temperature of the pad preceding the noise source and the cold insertion loss of the tube.

After taking all of these precautions, and adjusting the amplifier until everything was stable, a series of noise-figure measurements were made for frequencies very close to the degenerate frequency, $f_3/2$. The results of these measurements are shown in Table I, under the heading "Amplifier 2" along with theoretical values calculated from (34) to (36). The results in Table I for Amplifier 1 were obtained on an amplifier similar to the one described that had a higher cold insertion loss and less over-all gain. The measured values of F_V, the effective noise figures for wideband noise inputs are about 1.2 db higher than the calculated values for both amplifiers. This may be due to a variety of causes, some of which are:

1) Change in the insertion loss of the amplifiers with pumping due to an increase in the average value of the capacitances of the diodes.
2) Inaccurate calibration of the excess noise from the cold-cathode gas discharge noise source.
3) An accumulation of small errors due to the various sources of experimental error listed above.
4) Noise from the pump generator.
5) Small instabilities in the amplifier due to small changes in the pump power and signal level and even small thermal changes.

Note that the theoretical values of noise figures will decrease as the gain of the amplifiers increases. For Amplifier 2, with 3-db cold loss, F_V approaches 0.4 db and F_I approaches 3.4 db as the gain of the amplifier approaches infinity.

TABLE I
EXPERIMENTAL AND THEORETICAL NOISE FIGURES

Amplifier	Net Gain (Signal plus Idling)	Cold Loss	Measured Noise Figure F_V		Theoretical Noise Figures	
			Maximum Limits	Average	F_V (noise)	F_I (sig. gen.)
1	5.25 db	4.3 db	3.10–3.33 db	3.25 db	2.00 db	3.88 db
2	9.5 db	3.0 db	2.26–2.36 db	2.32 db	1.11 db	3.66 db

CONCLUSIONS

The traveling-wave parametric amplifiers described operate in the manner predicted theoretically, indicating that the techniques developed for their design are useful and valid. The particular type of amplifier built is limited in its usefulness by two principal factors: 1) the high gain at the edges of the pass band due to the high image impedance of the periodically loaded structure, and 2) the ripples in the pass band when high gain is required that are caused by small reflections from the amplifier terminations and some interaction with the backward waves. The introduction of nonreciprocal or even reciprocal loss can alleviate both of these problems. Alternative circuits with image impedances approaching zero at the edge of the band would eliminate the first difficulty. It may also be possible to synthesize nonperiodic filter structures which would avoid both of the above problems.

[28] S. B. Cohn and F. S. Coale, "Directional channel-separation filters," PROC. IRE, vol. 44, pp. 1018–1024; August, 1956.
[29] J. S. Honda and E. D. Sharp, "Investigation of Methods of Scanning the Beam of Large Antennas," Stanford Res. Inst., Menlo Park, Calif., Scientific Report 8, Contract AF 19(604)-2240; June, 1959.

On the Theory of Strongly Coupled Cavity Chains*

M. A. ALLEN† AND G. S. KINO†

Summary—A chain of identical cavity resonators coupled together through slots in their common walls forms a band-pass microwave filter. The pass band characteristics of such a system are determined by a combination of field theory and circuit theory. The fields in the cavities are expressed in terms of the normal modes of the uncoupled cavities. The fields in the neighborhood of a slot are determined by representing the slot as a transmission line. Irrotational components of the field in the cavities account for direct slot-to-slot coupling. The method successfully predicts both the dispersion characteristics and field distributions over large frequency ranges for many practical systems, such as slow-wave circuits for high-power traveling-wave tubes.

I. Introduction

THE theoretical determination of the dispersion relations and the field distributions of a system of coupled cavity resonators is, in general, a complicated problem. Analytic relationships for the resonant frequencies of a pair of coupled cavities have been obtained by Bethe[1] and have been applied to the case of weak coupling between cavities through small circular apertures. His analysis was facilitated by the assumption that the aperture was small compared to the wavelength, so that he was able to assume that the fields in the region of the hole have a static distribution. In this paper, strong coupling will be treated for a system in which one dimension of the coupling aperture is comparable in length to a wavelength. Chains of strongly-coupled cavity resonators are used in practice as slow-wave circuits for high-power traveling-wave tubes. In particular, the case of a chain of cavities coupled by long narrow slots in their common walls will be considered. A typical system with which we shall deal is shown in Fig. 1. For interaction with an electron beam, a beam-passage hole would be cut in the common walls. The amount of coupling introduced by such a hole is small compared with the coupling through the slots.

The case of coupling by long narrow slots has been treated in the past by assuming a quasi-static field distribution across the narrow dimension of a slot. In order to find the coupling through a slot cut in the common wall of a pair of waveguides, Stevenson[2] used this assumption and the further one of a sinusoidal voltage distribution along the slot; because of this latter assumption his analysis applies only near the slot resonance. Akhiezer and Lyubarsky,[3] who considered cavities coupled by long narrow slots, also employed the electrostatic assumption for a narrow slot, but derived the field distribution along the length of the slot by employing the theory of slot antennae. Each of the two preceding formulations is long, quite involved, and difficult to use in practice. In this paper a new method of approach will be given which makes use of transmission-line theory to determine the excitation of a slot in a connecting wall which can be of finite thickness. This theory is relatively simple to apply, and is, moreover, formulated in such a manner that the physical behavior of the system is always readily apparent.

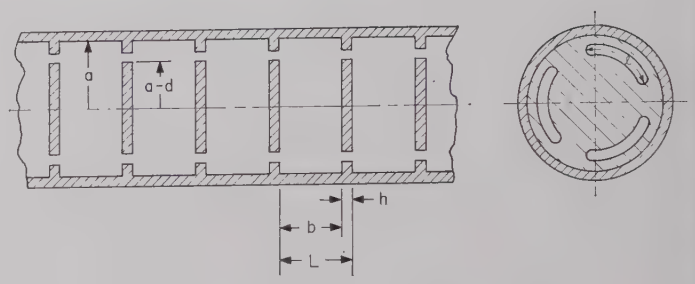


Fig. 1—Cylindrical cavity resonators coupled by long slots.

II. The Determination of the Dispersion Equation for a Chain of Identical Coupled Cavities

We shall consider in this paper a chain of identical loss-free symmetrical cavities coupled to each other through long slots cut in their common walls. We shall take the volume of the ith cavity of the system to be V_i ($V_i = V_{i+1}$, etc.) and shall denote by S_i' the area enclosed on the ith cavity wall by the slot or slots cut in the wall between the ith and $i-1$th cavities. Similarly, the area enclosed on the ith cavity wall by slots cut in the wall between the ith and $i+1$th cavities will be denoted by S_i''. The remaining surface area of the ith cavity will be denoted by S_i.

We shall express the fields in a cavity in terms of an expansion in the normal modes which could exist in the cavity in the absence of the slots; the coefficients of this expansion, or the excitation of such normal modes, will depend on the tangential component of electric field at the slots. The electric and magnetic fields of the nth normal mode of the ith cavity will be taken to be $\overline{E}_{i,n}$ and $\overline{H}_{i,n}$ respectively, the characteristic radian fre-

* Manuscript received by the PGMTT, December 1, 1959; revised manuscript received, January 28, 1960. This research was supported in part by the U. S. Air Force under Contract No. AF 30(602)-1844, monitored by the Rome Air Dev. Ctr. of the Air Res. and Dev. Command; and, in part, jointly by the U. S. Army Signal Corps, the U. S. Air Force, and the U. S. Navy (Office of Naval Res.).

† Microwave Labs., W. W. Hansen Labs. of Physics, Stanford University, Stanford, Calif.

[1] H. A. Bethe, "Theory of diffraction by small holes," *Phys. Rev.*, vol. 66, pp. 163–182; October, 1944.

[2] A. F. Stevenson, "Theory of slots in rectangular waveguides," *J. Appl. Phys.*, vol. 19, pp. 24–38; January, 1948.

[3] A. I. Akhiezer and G. Ya. Lyubarsky, "On the theory of coupled cavity resonators," *J. Tech. Phys. (USSR)*, vol. 24, pp. 1697–1708; September, 1954.

quency being $\omega_{i,n}$ with $\exp(j\omega_{i,n}t)$ dependence of all field quantities assumed.

The fields of these normal modes obey Maxwell's equations individually, and satisfy the boundary conditions at the perfectly conducting metal walls of the cavity, which are, for the ith cavity,

$$\bar{n}_i \times \bar{E}_{i,n} = 0 \quad \text{on} \quad S_i + S_i' + S_i''$$

and

$$\bar{n}_i \cdot \bar{H}_{i,n} = 0 \quad \text{on} \quad S_i + S_i' + S_i' \tag{1}$$

where $\bar{n}_i$ is the unit normal vector at the surface of the ith cavity pointing into the ith cavity.

It has been shown by Teichmann and Wigner[4] that the normal modes do not form a complete set of functions in terms of which any possible fields in the system of coupled cavities can be expressed. In the presence of coupling holes and loops, it is necessary to add irrotational components of electric and magnetic field to form the complete expansion. When only coupling slots are considered to be present, Teichmann and Wigner have shown that only an irrotational magnetic field, in addition to the normal mode fields, is required. In order to point out the principles of our method of approach, we shall first set up the complete expansion, but in the first application of the method we shall neglect the contribution of this irrotational part of the magnetic field. However, in a later part of the paper, we shall show that it is often necessary to consider this component of the field; it is, in fact, extremely important when the periodic length of the system is very much less than a slot width, since it accounts for the direct coupling from slot to slot of a multicavity chain and changes the effective slot resonant frequency.

We shall denote the irrotational component of the magnetic field in the ith cavity by $\bar{H}_{i,0}$. By definition

$$\nabla \times \bar{H}_{i,0} = 0. \tag{2}$$

Hence, $\bar{H}_{i,0}$ can be expressed as the gradient of a scalar quantity ψ_i by writing

$$\bar{H}_{i\,0} = \nabla\psi_i. \tag{3}$$

Using an approach similar to that of Slater,[5] as shown in Appendix A, the fields $\bar{E}_i$ and $\bar{H}_i$ that exist in the ith cavity of the coupled system may be expressed as follows:

$$\bar{E}_i = \sum_n e_{i,n}\bar{E}_{i,n} = \sum_n \left[\frac{j\omega_{i,n}}{\omega^2 - \omega_{i,n}^2} \cdot \frac{\int_{S_i'+S_i''} (\bar{E}_i \times \bar{H}_{i,n}^*)\cdot \bar{n}_i dS}{2W_{i,n}} \bar{E}_{i,n} \right] \tag{4}$$

$$\bar{H}_i = \sum_n h_{i,n}\bar{H}_{i,n} + \nabla\psi_i = \sum_n \frac{j\omega}{\omega^2 - \omega_{i,n}^2} \cdot \left[\frac{\int_{S_i'+S_i''} (\bar{E}_i \times \bar{H}_{i,n}^*)\cdot \bar{n}_i dS}{2W_{i,n}} \bar{H}_{i,n} \right] + \nabla\psi_i, \tag{5}$$

where

$$W_{i,n} = \frac{1}{2}\int_{V_i} \epsilon E_{i,n}^2 d\tau = \frac{1}{2}\int_{V_i} \mu H_{i,n}^2 d\tau. \tag{6}$$

In the ith cavity, because $\nabla\cdot\bar{H}_{i,0}=0$, we have the condition

$$\nabla^2\psi_i = 0 \tag{7}$$

with the boundary condition

$$\bar{n}_i\cdot\nabla\psi_i = \bar{n}_i\cdot\bar{H}_i \quad \text{on} \quad S_i' + S_i''. \tag{8}$$

On the surface of a connecting slot, the normal component of $\bar{H}$ is not zero, so that $\bar{H}_{i,0}$ exists in the ith cavity, and it is necessary to include it in the normal mode expansion.

Thus the fields in the ith cavity have been expressed entirely in terms of the tangential component of the electric field and the normal component of magnetic field at the slots which are cut in its walls.

We determine the field distribution along a slot by considering it to consist of a pair of finite-width parallel conductors with perfect shorts placed across them at planes corresponding to the ends of the slot. This is illustrated in Fig. 2. Because the slot width, or conductor spacing, is assumed to be much smaller than a wavelength, it is reasonable to assume that the electric field lines in the neighborhood of the slot always lie in a plane perpendicular to the longitudinal dimension of the slot; thus there can only be pure TE waves or TEM waves existing in the region of the slot. For a start, we shall assume that the slot fields themselves are pure TEM; this assumption may be shown to be equivalent to neglecting the presence of the irrotational mode. It is then possible to express the electrical properties of the slot in terms of transmission-line theory, the inductance and capacitance of the transmission line being computed by static considerations from the transverse dimensions of the slot. In order to write down suitable transmission-line equations to represent the excitation of a slot by the cavity fields, it is convenient to divide the currents flowing in the neighborhood of the slot into two parts. The first part is the current, I, which flows around the slot or along the length of the transmission line; this current is directly associated with the component of magnetic field normal to the plane of the slot. The second part is the exciting current, j_y, which flows on the surface of a cavity into the slot in a direction perpendicular to the longitudinal dimension of the slot; this current is associated with the component of magnetic field in the cavities which lies along the longi-

[4] T. Teichmann and E. P. Wigner, "Electromagnetic field expansions in loss-free cavities through holes," *J. Appl. Phys.*, vol. 24, pp. 262–267; March, 1953.

[5] J. C. Slater, "Microwave Electronics," D. Van Nostrand Co., Inc., New York, N. Y.; 1950.

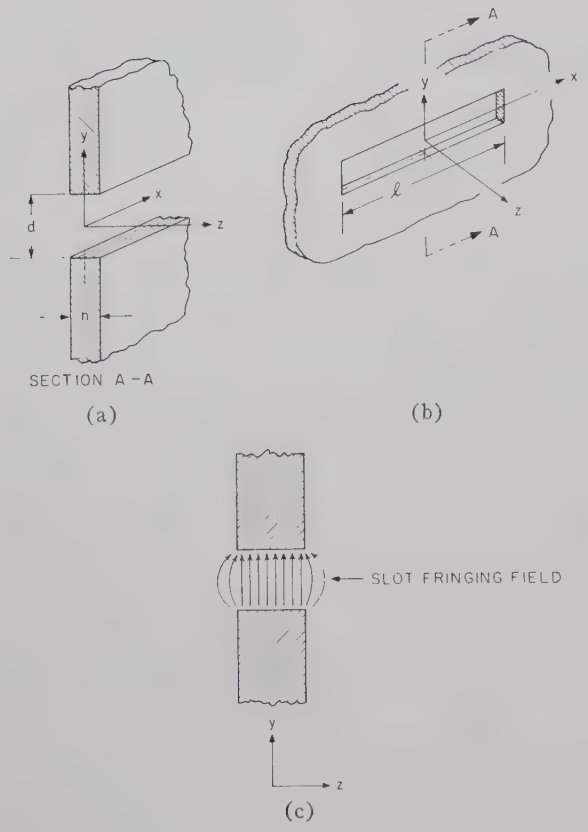


Fig. 2—(a) and (b) Views of slot with coordinate system shown. (c) Section through slot showing behavior of the electric field.

tudinal dimension of the slot. For a slot which is cut in the common wall of the $i-1$th and ith cavities, the current j_y flowing into the slot is given in value by the expression

$$j_y = j_{i-1,y}'' + j_{i,y}' = \bar{n}_i \times (\overline{H}_i + \overline{H}_{i-1})_x \text{ on } S_i' + S_{i-1}'' \quad (9)$$

where the subscript y represents the direction along the surface of a cavity, perpendicular to the longitudinal direction of a slot and j_y denotes the *total* current flowing into the slot on both sides of the wall between the $i-1$th and ith cavities. The significance of these two current components is illustrated in the transmission-line representation in Fig. 3. This representation yields two differential equations for the voltage and current distributions along a slot in terms of the fields at a slot; these may be written as follows:

$$\frac{\partial \phi}{\partial x} = -j\omega L I \quad (10)$$

$$\frac{\partial I}{\partial x} - j_y = -j\omega C \phi \quad (11)$$

with the boundary conditions

$$\phi = 0 \quad \text{at} \quad x = \pm \frac{l}{2}, \quad (12)$$

where C is the capacitance per unit length of the slot, L is the inductance per unit length of the slot, ϕ the voltage across a slot, l the length of a slot, and x denotes distance along the slot from its center. Eqs. (10) and (11) yield the differential equation for the voltage distribution across a slot, in terms of the driving currents due to the cavity modes,

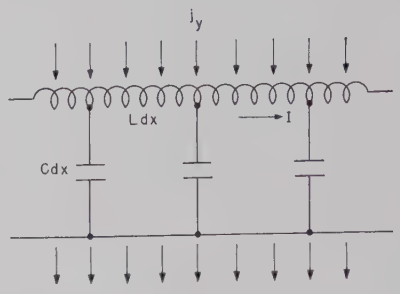


Fig. 3—Transmission-line representation of the slot with net current flow across the slot.

$$\frac{\partial^2 \phi}{\partial x^2} + k^2 \phi = -jkZ_0 j_y \quad (13)$$

where $k = \omega/c$ and $Z_0 = \sqrt{L/C}$, the characteristic impedance of the slot. Writing this equation as

$$\frac{\partial^2 \phi}{\partial x^2} + k^2 \phi = -f(x), \quad (14)$$

the solution (see, for example, Webster[6]) may be expressed as follows:

$$\phi = \frac{\sin k(l/2 - x)}{k \sin kl} \int_{-l/2}^{x} \sin k(l/2 + \xi) f(l/2 + \xi) d\xi + \frac{\sin k(x + l/2)}{k \sin kl} \int_{x}^{l/2} \sin k(l/2 - \xi) f(l/2 - \xi) d\xi. \quad (15)$$

By writing

$$\phi = \int_0^d E_y dy, \quad (16)$$

we thus have an expression for the tangential component of the electric field at the slot in terms of the cavity fields in the neighborhood of a slot of width d. The characteristic impedance Z_0 can be determined analytically by a conformal transformation or experimentally in an electrolytic tank. Thus finite connecting wall thickness may be taken into account. For the limiting case of zero wall thickness the inverse cosine transformation may be used.[7]

We have in (4) and (5) an expression for the fields in the ith cavity in terms of the tangential component of the electric field at (*i.e.*, the voltage across) the slots. The current which flows into the region of a slot is expressible in terms of these fields and hence in terms of the voltages across the slots. This current is the current

[6] A. G. Webster, "Partial Differential Equations of Mathematical; Physics," Dover Publications, New York, N. Y.; 1955.

[7] S. Ramo and J. R. Whinnery, "Fields and Waves in Modern Radio," John Wiley and Sons, Inc., New York, N. Y., 2nd ed., pp. 135–138; 1953.

which excites the slot fields, and in terms of which the voltage across a slot was expressed in (13) and (15). As a result, we have an expression for the current flowing into the region of a slot in terms of the voltage across a slot, and an expression for the voltage across a slot in terms of the current flowing into the slot region. On combining these expressions, a single integral equation whose eigenvalues give the dispersion equation of the coupled system may be obtained. For simplicity, we shall only consider a system made up of identical cavities coupled by a single slot in each cavity wall; the extrapolation of the method to a multislot coupling system may be readily accomplished.

Following the procedure outlined above, we first write down the current j_y in terms of the normal mode fields in the ith and $i-1$th cavities. Thus from (5) and (9) it follows that

$$j_y = \bar{n}_i \times \Big[\sum_n (h_{i,n}\overline{H}_{i,n})_{S_i'} + \sum_n (h_{i-1,n}\overline{H}_{i-1,n})_{S_{i-1}''} \Big]_y \quad (17)$$

where, at this juncture, we have neglected the contribution of the irrotational field. A further transformation of (17), by writing it in the form

$$j_y = \sum_n (h_{i,n} j_{i,n,y}') + \sum_n (h_{i-1,n} j_{i-1,n,y}'') \quad (18)$$

is convenient, where $j_{i,n}'$ and $j_{i-1,n}''$ are the surface currents, in the region of the slot of the nth normal modes in the ith and $i-1$th cavities respectively. In terms of the normal mode fields

$$j_{i,n}' = (\bar{n}_i \times \overline{H}_{i,n})_{S_{i,n}'}$$
$$j_{i-1,n}'' = (\bar{n}_i \times \overline{H}_{i-1,n})_{S_{i-1,n}''}. \quad (19)$$

The form of these equations may now be simplified by taking the nth normal mode fields in each cavity to be identical, so that we can use Floquet's theorem to write

$$h_{i-1,n} = h_{i,n} e^{j\beta L} \quad (20)$$

where L is the periodic length of the system, and βL the phase delay between cavities.

Moreover, as the cavities are symmetrical and identical,

$$j_{i,n}'' = j_{i-1,n}'' = \pm j_{i,n}' \quad (21)$$

where the choice of sign depends on the form of the nth mode of the cavity. For example, the negative sign is appropriate for the TM_{010} mode of a cylindrical cavity, the positive sign for the TM_{011} mode of the same cavity. Thus, we can simplify (18) by the use of (20) and (21) and write

$$j_y = \sum_n h_{i,n}(1 \pm e^{j\beta L}) j_{i,n,y}'. \quad (22)$$

In order to determine the final dispersion relation, we must first find $h_{i,n}$ in terms of j_y. To do this, we may make use of (5) and (19) to write

$$h_{i,n} = \frac{-j\omega}{\omega^2 - \omega_{i,n}^2} \cdot \frac{\int_{S_i''} \overline{E}_i \cdot \bar{j}_{i,n}'' dS + \int_{S_i'} \overline{E}_i \cdot j_{i,n}' dS}{2W_{i,n}}. \quad (23)$$

We may use Floquet's theorem and (21) to write (23) in the form:

$$h_{i,n} = \frac{-j\omega}{\omega^2 - \omega_{i,n}^2} \cdot \frac{\int_{S_i''} \overline{E}_i \cdot \bar{j}_{i,n}' (1 \pm e^{-j\beta L}) dS}{2W_{i,n}}. \quad (24)$$

In order to express (24) in terms of the voltage across the slot, it is convenient to make the assumption that $j_{i,n}'$ is constant over the width of the slot, although this assumption is not strictly necessary. With this assumption, however, we may carry out the integration over the width of the slot and write

$$h_{i,n} = \frac{-j\omega}{\omega^2 - \omega_{i,n}^2} \cdot \frac{(1 \pm e^{-j\beta L}) \int_{-l/2}^{l/2} \phi_i j_{i,n,y}' dx}{2W_{i,n}}. \quad (25)$$

We now find j_y in terms of ϕ by substituting (25) in (22) to yield the following expression:

$$j_y = \sum_n \frac{-j\omega(1 \pm \cos\beta L) j_{i,n,y}' \int_{-l/2}^{l/2} \phi_i j_{i,n,y}' dx}{(\omega^2 - \omega_{i,n}^2) W_{i,n}}. \quad (26)$$

It will be recalled that we have already found ϕ_i in terms of j_y in (15); thus we are in a position to eliminate ϕ_i from (15) and (26) and to write the following integral equation for j_y:

$$j_y = \sum_n \frac{-j\omega(1 \pm \cos\beta L) j_{i,n,y}'}{(\omega^2 - \omega_{i,n}^2) W_{i,n}} \int_{-l/2}^{l/2} dx \cdot \Big\{ j_{i,n,y} \Big[\frac{\sin k(l/2 - x)}{k \sin kl} \int_{-l/2}^{x} \sin k(l/2 + x) f(l/2 + \xi) d\xi + \frac{\sin k(x + l/2)}{k \sin kl} \int_{x}^{l/2} \sin k(l/2 - \xi) f(l/2 - \xi) d\xi \Big] \Big\} \quad (27)$$

where

$$f(x) = -jkZ_0 j_y(x).$$

Eq. (27) is a Fredholm integral equation of the second kind which may in principle be solved to give the possible modes of the coupled system. Such a procedure is, in practice, too difficult, so that it is necessary to make certain approximations of a type which will be described in the next section. Alternatively, it is sometimes convenient and more accurate to set up the solution for this coupled system in a variational form. The procedure for obtaining this variational form of the solution is given in Appendix B.

III. Application to a Particular Cavity Chain

The particular resonators which we shall study for the purpose of illustrating our methods will be cylindrical cavities coupled to one another by long, circumferential, narrow, identical slots cut in the common walls of adjacent cavities, as shown in Fig. 1.

The dominant cavity mode in the system, before the slots are cut in the disks, will be taken to be the TM_{010} mode. Its characteristic radian frequency will be denoted by ω_1.

We have developed an expression, (27), for the characteristic frequencies of a coupled system. We shall make use of an approximation to this expression which employs only the dominant (TM_{010}) mode in the series expansion. This gives rise to a simplified coupling equation. At a later stage, the effect of the irrotational magnetic mode introduced earlier will be considered.

A. Dominant Mode Expansion

The circumferential coupling slot will be taken to be cut close to the cylindrical walls. The slot length and thickness are denoted by l and h, respectively, and the cavity radius and length by a and b, respectively. Thus, the periodic length of the system is $L=a+h$. The dominant TM_{010} mode of the cavities has the following fields:

$$E_z = E_0 J_0(k_1 r)$$
$$H_\phi = \frac{jE_0}{\eta} J_1(k_1 r) \tag{28}$$

where $J_\nu(x)$ is a Bessel function of the first kind and νth order; and

$$k_1 = \frac{p_{01}}{a} = \frac{2.405}{a},$$

p_{01} being the first zero of the zero-order Bessel function. We shall refer to this dominant mode by the subscript 1. In the normal mode expansion we shall only use this dominant mode term to determine the coupling. The fields in the cavities are therefore expressed as follows:

$$\overline{E}_i = e_{i,1}\overline{E}_{i,1} + \overline{E}_i'$$
$$\overline{H}_i = h_{i,1}\overline{H}_{i,1} + \overline{H}_i', \tag{29}$$

where $\overline{E}_i'$ and $\overline{H}_i'$ are the remaining fields orthogonal to the dominant mode. The current j_y which drives the slot is assumed to arise only from the dominant mode fields. These fields have no θ variation, which, for circumferential slots, means that they have no x variation. With the current being independent of the x coordinate, (15) gives ϕ_i in the form

$$\phi_i = -\frac{jZ_0 j_y}{k}\left(1 - \frac{\cos kx}{\cos kl/2}\right). \tag{30}$$

By retaining only the dominant mode in (27), and noting that j_y is now independent of the x coordinate, we obtain the following dispersion relation:

$$\frac{\Omega_1(1-\Omega_1^2)}{\tan\left(\frac{\rho\Omega_1\pi}{2}\right) - \frac{\rho\Omega_1\pi}{2}} = \alpha_\pi \sin^2\frac{\beta L}{2} = \alpha \tag{31}$$

where

$$\alpha_\pi = \frac{4j_1^2 Z_0 c^2}{W_1 \omega_1^2}; \tag{32}$$

and

$$\Omega_1 = \frac{\omega}{\omega_1}; \qquad \rho = \frac{\omega_1}{\omega_{s1}}; \qquad \frac{\omega_{s1} l}{2c} = \frac{\pi}{2}.$$

Since the cavities are identical we now omit the subscript referring to the particular cavity.

Eq. (31) yields on solution an infinite number of values of ω for each value of α, the dimensionless coupling coefficient. For $\alpha_\pi \sin^2 \beta L/2 = 0$ (for this system, when $\beta L = 0$) these values of ω are ω_1, ω_{s1}, ω_{s2}, ω_{s3}, $\cdots$, corresponding to the poles of the denominator and zero of the numerator of (31). Solving the equation as βL increases in value from zero to π, an infinite number of pass bands results, one associated with the cavity dominant mode characteristic frequency ω_1, and the others associated with the frequency ω_{sp} which makes the slot p half-wavelengths long. We shall concern ourselves only with the two lowest frequency pass bands: the one associated with ω_1, and the other associated with ω_{s1}. Plotted on an ω-β diagram, the higher frequency curve has a positive slope ($d\omega/d\beta > 0$) and the low frequency curve has a negative slope ($d\omega/d\beta < 0$).

Numerical solutions of (31) are given, with the quantity Ω_1 plotted as a function of ρ, from $\rho = 0.5$ to $\rho = 2.5$ for several values of α from zero to unity in Fig. 4. With α known, the ω-β diagram may be plotted for the coupled-cavity system by taking values from Fig. 4, using the appropriate value of ω_1/ω_{s1}.

In order to determine the dispersion relation for a particular system, we must evaluate the coupling coefficient. In the dominant mode we have

$$j_{1y}^2 = \frac{E_1^2(0)}{\eta^2} J_1^2(k_1 a), \tag{33}$$

assuming the slot is cut near the maximum radius where the magnetic field is a maximum.

Since the dominant mode has no θ variation, the energy stored in this mode, W_1, is expressed by a single integral,

$$W_1 = b\int_0^a \frac{\epsilon |E_z|^2 2\pi r dr}{2}. \tag{34}$$

On substituting (33) and (34) into (32) we obtain

$$\alpha_\pi = \frac{8Z_0 c^2}{\eta^2 \pi \epsilon b a^2 \omega_1^3}. \tag{35}$$

For the TM_{010} mode $\lambda_1 = 2.61a$; hence

$$\omega_1^3 = \frac{8\pi^3 c^3}{17.8a^3}, \tag{36}$$

and since $\epsilon = 1/\eta c$ where η is the characteristic impedance of free space, we obtain finally a simple formula for the coupling coefficient

$$\alpha_\pi = 0.18\left(\frac{a}{b}\right)\left(\frac{Z_0}{\eta}\right). \tag{37}$$

This coupling coefficient, as we might expect, depends upon the static properties of the slot, Z_0, and the ratio of the cavity radius to the cavity length, a/b.

The dispersion characteristics for a typical structure are shown in Fig. 5. Agreement between theory and experiment is seen to be good in the cavity pass band. In the slot pass band some disagreement is observed, and in structures where the slot width is comparable to the cavity length ($d \approx b$ in Fig. 1) this disagreement becomes large. This is a result of the exclusion of the irrotational component of the magnetic field from the field expansion.

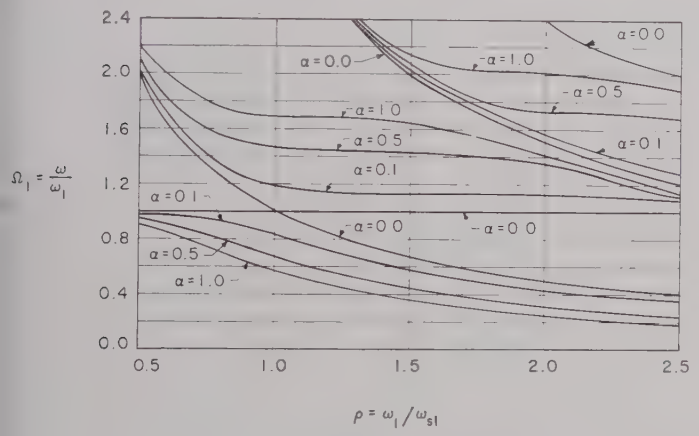


Fig. 4—Solutions to the dispersion relation (31).

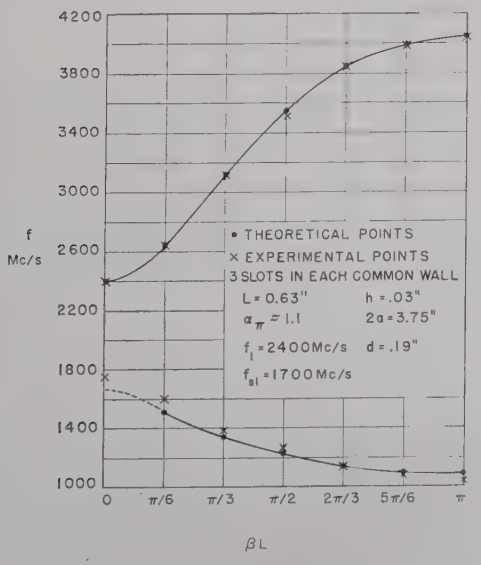


Fig. 5—Typical Brillouin diagram between $\beta L = 0$ and $\beta L = \pi$ for the two lowest pass bands for a coupled-cavity system of the type shown in Fig. 1.

B. *Inclusion of the Irrotational Mode*

We shall now consider the way in which the irrotational magnetic mode arises in the coupled system under study. It was shown that the fields in the slots gave rise to voltages and currents which obey the normal transmission-line equations with an additional driving current term. The current which flows along the slot is directly related to the normal component of magnetic field at the slot. If this slot were cut in an infinite plane, it could be assumed that the slot fields were pure TEM. However, because the slot is actually cut in the wall of a closed resonator, the magnetic field lines must be continuous through the slot and must be closed, as shown in Fig. 6. Therefore, a longitudinal component of the magnetic field must exist which is associated with the normal magnetic field at the slot. Mathematically, this longitudinal field component can only be expressed by the irrotational field.

This longitudinal component of magnetic field, which is represented by the irrotational mode, has two effects: 1) It gives the effect of shunt inductance across the slot and, hence, changes the resonant frequency of the slot. 2) It introduces a component of field which can directly provide large amounts of slot-slot coupling or mutual inductance between the slots. In order to illustrate what occurs, we consider the system of slot-coupled resonators shown in Fig. 7. There will be an irrotational component of magnetic field originating from the currents which flow along slot A. Associated with this irrotational magnetic field there will be currents which flow in the walls of the two cavities perpendicular to the slot A. The wall current which flows

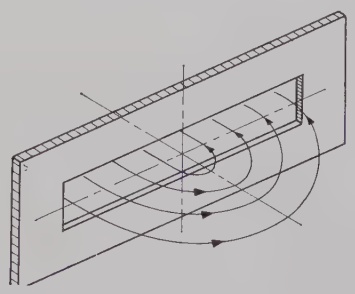


Fig. 6—Sketch of magnetic field lines associated with the normal component of the magnetic field at the slot.

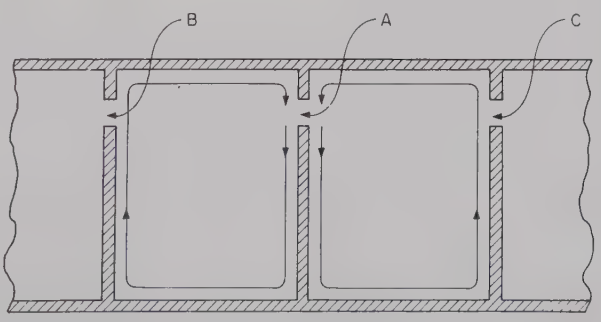


Fig. 7—Current associated with the longitudinal component of the irrotational part of the magnetic field at the slot.

across slot A we call the "self" current; it may be represented by the current flowing in a shunt inductance across slot A, and it causes an increase in the resonant frequency of the slot. Similar currents arise from the slots B and C adjacent to slot A in Fig. 7. A portion of these currents flow into slot A as "mutual" currents, giving slot-slot coupling which may be represented as mutual inductance between the adjacent slots A-B and A-C. If there is zero phase shift between adjacent slots, the mutual current from the two slots adjacent to a given slot and the self-current of the slot will tend to cancel. For a $\pi/2$ phase shift there will be no net effect from the mutual currents arising from adjacent slots, and at π phase shift the self and mutual currents will add. The over-all effect is that the bandwidth of the slot band is reduced below the value predicted by a theory which does not include the irrotational component of the magnetic field. The effect of this extra component of field is less important in the cavity pass band than in the slot pass band.

The situation which has been described is that which is the case for identically placed slots between the cavities. If the slots at one plane are rotated with respect to the slots at adjacent planes the effect of the mutual currents is reduced.

Fig. 8 gives experimental results in a system with the slots aligned and in a system with the slot rotated 90° to adjacent slots. It is noted that with the mutual effect eliminated, the self-effect raises the slot band in frequency but does not reduce its total bandwidth.

The behavior of this irrotational mode depends on the determination of the function which is the solution to the Neumann problem in a cylindrical cavity resonator with the given boundary conditions. With the problem solved, the gradient of this function is taken along the slot itself and along adjacent slots to give the self- and mutual currents, respectively. The solving of the Neumann problem involves dealing with slowly converging series of trigonometric and Bessel functions. The convergence of these series depends critically on the fact that the width of the slot is finite. It has not been possible to sum these series in closed form; consequently only a semiquantitative picture will be given of the effect of the irrotational mode, or the "slot-slot" coupling which it represents. The slot is of width d, and is placed at $z=L$, as shown in Fig. 9. We have for a given cavity

$$\nabla^2\psi = 0$$

$$\frac{\partial\psi}{\partial n} = H_n \quad \text{on} \quad S' + S''$$

$$\frac{\partial\psi}{\partial n} = 0 \quad \text{on} \quad S. \tag{38}$$

The most general solution for ψ which obeys Laplace's equation and the boundary conditions of the cylindrical cavity is

$$\psi = \sum_{mn} A_{mn} \cosh(\beta_n z) J_m(\beta_n r) \sin(m\theta) \tag{39}$$

with

$$J_m'(\beta_n a) = 0.$$

The current along the slot, I, is related to the normal magnetic field H_n at the slot as follows:

$$LI = \mu H_n d \tag{40}$$

where L is the inductance per unit length of the slot, as defined previously. The field is assumed to be constant across the width of the slot. From (10), (11), and (30) we have

$$I = I_0 \sin(ak\theta), \tag{41}$$

where I_0 is the maximum value of the total current flowing along the slot. We note that $L/\mu = Z_0/\eta$ where Z_0 and η are the characteristic impedances of the slot and free space, respectively. Thus, from (40), at the slot we have

$$\left.\frac{\partial\psi}{\partial z}\right|_{z=L} = \frac{Z_0}{\eta}\frac{I_0}{d}\sin(ak\theta). \tag{42}$$

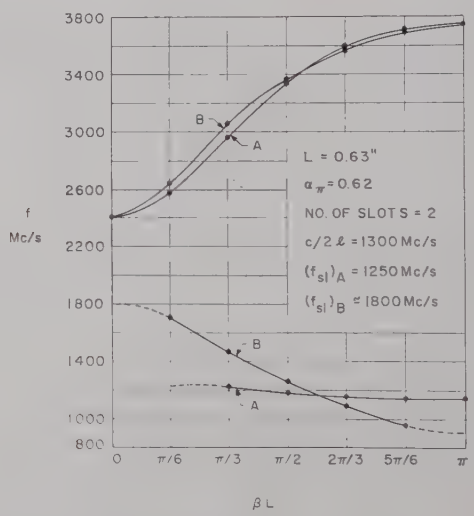


Fig. 8—Brillouin diagrams for coupled-cavity system. A. Adjacent slots in line. B. Adjacent slots rotated 90°.

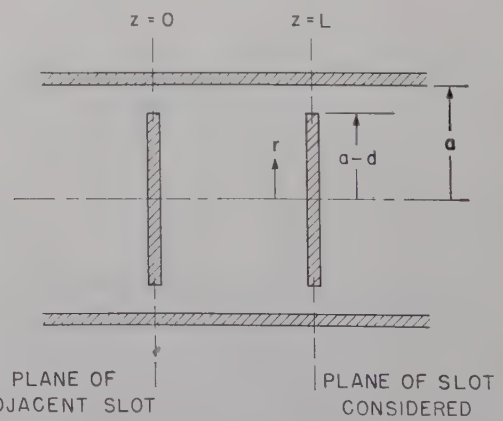


Fig. 9—Coordinate system for calculation of irrotational component of the magnetic field.

Eqs. (38), (39), and (42) may be combined[8] to give an approximate expression for the irrotational field in the cavities and the self- and mutual currents associated with this field give an added current flowing in the slot region at the cavity

$$j_{y0} = -P\frac{\partial I}{\partial x} + Q\frac{\partial I}{\partial x}\cos(\beta L), \tag{43}$$

where

$$P = \frac{Z_0}{\eta}\sum_n \frac{1}{\tan(\beta_n L)}\frac{\beta_n a}{(\beta_n a)^2 - 1}\left(\frac{\sin(\beta_n d/2)}{\beta_n d/2}\right)^2$$

$$Q = \frac{Z_0}{\eta}\sum_n \frac{1}{\sinh(\beta_n L)}\frac{\beta_n a}{(\beta_n a)^2 - 1}\left(\frac{\sin(\beta_n d/2)}{\beta_n d/2}\right)^2. \tag{44}$$

We have made the approximation of using only the first θ term in the double summation of (39). The strongly converging terms

$$\left(\frac{\sin(\beta_n d/2)}{\beta_n d/2}\right)^2$$

arise from integrations over the finite width of the slots. The net current driving the slot is $j_y + j_{y0}$. Modifying (11) accordingly, (31) now becomes

$$\frac{\Omega_1(1-\Omega_1^2)}{\tan(\rho'\Omega_1\pi/2) - \rho'\Omega_1\pi/2} = \alpha\sqrt{1+P-Q\cos\beta L}\,\sin^2(\beta L/2) \tag{45}$$

where

$$\rho' = \frac{\rho}{\sqrt{1+P-Q\cos\beta L}}.$$

Thus, we have a coupling equation which is of the same form as the coupling equation obtained without the inclusion of the irrotational mode, but with different slot resonant frequencies and coupling coefficients,

$$\omega_{s1}' = \omega_{s1}\sqrt{1+P-Q\cos\beta L}$$

$$\alpha_\pi' = \alpha_\pi\sqrt{1+P-Q\cos\beta L}. \tag{46}$$

The quantity P, the self-term, will in general be larger than the quantity Q, the mutual term, from adjacent slots. The slot resonant frequency will be higher than in the case without the inclusion of the irrotational mode at $\beta L = 0$. As βL increases, the slot resonant frequency will increase in value. Based on this higher slot resonant frequency, the coupling between cavities will, for certain βL, give a characteristic frequency below this resonant frequency, as is given by the equation whose solutions are given in Fig. 4. The net result will be a decrease in the total bandwidth of the pass band associated with the slot resonant frequency. The sum of the first ten terms of the expansion, in (44), gives a good approximation of the expansions. In (45) the values of P and Q thus obtained account for the departure from theory in the lower pass band of Fig. 5.

IV. Determination of Field Distributions

In assessing the usefulness of microwave filters for interaction with an electron beam, the quantity $E^2(0)/W$ is a figure of merit where $E^2(0)$ is the square of the longitudinal component of the electric field on the axis and W is the energy stored per periodic length. This quantity will be considered.

It is possible to make accurate estimates of the ω-β curves for large coupling between cavities by using a theory based on only the dominant mode in the normal mode expansions of the fields. Although the ω-β curves can be accurate at frequencies far from the dominant mode resonance, the dominant mode alone is not sufficient to express the values of the fields in the cavities. This is illustrated by the fact that, in Appendix B, we found an expression for ω in a variational form; thus, by using only a rough estimate of the fields in the slot we obtain a value for the frequency which is less in error than the estimate of the fields themselves.

A theorem exists for lossless periodic transmission system which states that the time-average electric stored energy per period is equal to the time-average magnetic stored energy per period in the pass band. A proof of this theorem has been given by several authors.[9] Using this theorem we are able to express the stored energy per period in terms of only the electric energy.

The cylindrical cavity resonators with which we are concerned have lengths much smaller than their diameters. The characteristic frequencies of the normal modes having variations of the field in the direction of the cavity length (z-varying modes) will thus be very much above the characteristic frequencies of the corresponding non-z-varying modes, and will have much smaller amplitudes in the expansion. Therefore, the electric energy stored in the coupled system under consideration, except in and around the region of the slot, comes largely from the dominant mode and other non-θ-, non-z-varying modes. The dominant mode has a longitudinal component of the electric field which has a maximum value on the axis and falls to zero at the circumference of the cavity. Higher-order modes have further zeros between the maximum value on the axis and zero at the circumference. In the coupled system, for values of frequency above the dominant mode frequency, since one or more additional zeros in the longitudinal component of the electric field must occur, the dominant mode alone cannot give a good representation of the field. This is illustrated in Fig. 10. Thus the stored energy in the cavities must have appreciable contributions from several modes.

[8] M. A. Allen, "Coupling of Multiple Cavity Systems," Microwave Lab., W. W. Hansen Labs. of Physics, Stanford University, Stanford, Calif., M.L. Rept. No. 584; 1959.

[9] D. A. Watkins, "Topics in Electromagnetic Theory," John Wiley and Sons, Inc., New York, N. Y.; 1958.

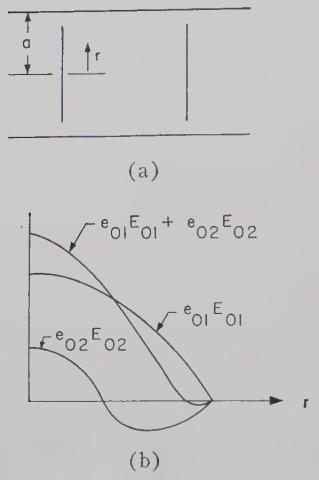


Fig. 10—For structure in (a), the plot of the amplitudes as a function of radial distance for the first two terms of the normal mode expansions is given in (b).

We now consider the slot region. If the total field $\bar{E}$ everywhere is divided into two parts

$$\bar{E} = \left(\sum_n e_{0n}\bar{E}_{0n} \right) + \bar{E}' \tag{47}$$

where the subscript $0n$ refers to the TM_{0n0} modes, then a large proportion of the electric energy stored in the slot region would arise from the part of the field represented by $\bar{E}'$. Because of the orthogonality of the normal modes, the stored energy W per periodic length is given by

$$\begin{aligned} W &= 1/2 \int_{V_i} \epsilon \bar{E} \cdot \bar{E}^* d\tau \\ &= \sum_n e_{0n}^2 W_{0n} + 1/2 \int_{V_i} \epsilon \bar{E}' \cdot \bar{E}'^* d\tau. \end{aligned} \tag{48}$$

The summation in (48) may be summed in closed form.[8] The volume integral on the right-hand side of (48) has most of its contribution from the part of the volume near the slot. The behavior of the field in the slot region was derived from considerations of the excitations of a TEM wave guided by parallel-plane conductors leading to a static description of the fields there. The energy in the fields at the slot is as would be derived for a static distribution of the fields. Thus, the time-average electric energy stored by a slot W_{sE} is given by

$$\begin{aligned} W_{sE} &= 1/2 \int_{V_i} \epsilon \bar{E}' \cdot \bar{E}'^* d\tau \\ &= 1/2 \int_{-l/2}^{l/2} C\phi^2 dx. \end{aligned} \tag{49}$$

We assume that (49) includes all the energy stored outside the TM_{0n0} modes in the coupled system, and so avoid having to consider the z-varying and θ-varying modes' contributions separately. The voltage ϕ may therefore be obtained from (30).

If now the field on the axis is considered, then, of the normal modes which have a longitudinal component of electric field on the axis, the non-z-varying modes will, as in the considerations of energy, contribute more to the normal mode expansion for the electric field on the axis than the z-varying modes. Thus, for cavities of large diameter-to-length ratio, we express the electric field on the axis as an expansion based only on the TM_{0n0} normal modes,

$$\bar{E}(0) = \sum e_{0n}\bar{E}_{0n}(0). \tag{50}$$

We note that the azimuthal-varying modes have no longitudinal component of the electric field on the axis, and they need not be considered as they are in computing W, the total stored energy per period. Using values of frequency based on the dominant mode expansion, (50) may be summed in a closed form.[8]

For the type of coupling we are considering, the irrotational mode makes no direct contribution to the electric energy stored per period; it contributes only to the stored magnetic energy. However, this irrotational mode does have an effect on the values of the coefficients of the normal mode expansions. Values of frequency based on this irrotational mode should be included in the determination of the quantity $E^2(0)/W$, especially in the slot band.

Some typical theoretical and experimental results are shown in Fig. 11. If only the dominant mode had been used in the normal mode expressions the large values of $E^2(0)/W$ at large values of βL in the upper pass band would not have been predicted.

Appendix A

Evaluation of the Coefficients of the Normal Mode Expansions of the Fields in a Cavity

The normal modes of the ith cavity are given by

$$\bar{E}_{i,n}(\bar{r}), \qquad \bar{H}_{i,n}(\bar{r}).$$

$\omega_n^{(i)}$ are the associated characteristic frequencies, with

$$e^{-j\omega_n^{(i)}t}$$

time dependence assumed. Maxwell's equations with electric boundary conditions are obeyed:

$$\left.\begin{aligned} \bar{\nabla} \times \bar{E}_{i,n} &= -j\omega_{i,n}\mu H_{i,n} \\ \bar{\nabla} \times \bar{H}_{i,n} &= j\omega_{i,n}\epsilon\bar{E}_{i,n} \\ \bar{\nabla} \times \bar{E}^*_{i,n} &= j\omega_{i,n}\mu\bar{H}^*_{i,n} \\ \bar{\nabla} \times \bar{H}^*_{i,n} &= -j\omega_{i,n}\epsilon\bar{E}^*_{i,n} \end{aligned}\right\} \tag{51}$$

with

$$\bar{n} \times \bar{E}_{i,n} = 0 \quad \text{on} \quad S_i + S_i' + S_i''. \tag{52}$$

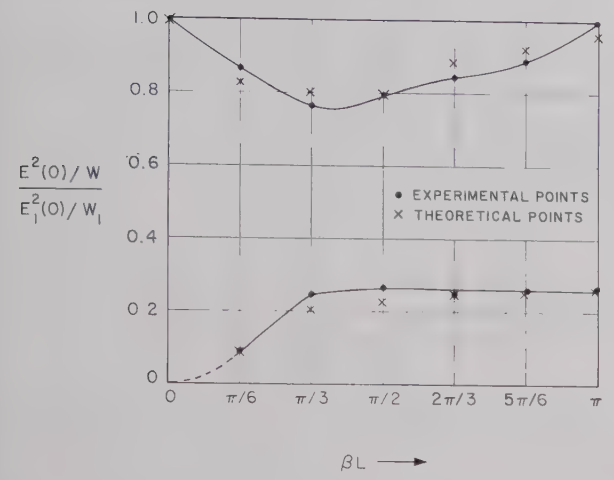


Fig. 11—Normalized $E^2(0)/W$ plotted as a function of βL for the two lower pass bands of a coupled-cavity system of the type shown in Fig. 1.

The required fields $\bar{E}(\bar{r})$ and $\bar{H}_i(\bar{r})$ satisfy Maxwell's equations with frequency ω in the coupled system:

$$\left.\begin{aligned} \bar{\nabla} \times \bar{E}_i &= -j\omega\mu\bar{H}_i \\ \bar{\nabla} \times \bar{H}_i &= j\omega\epsilon\bar{E}_i \\ \bar{\nabla} \times \bar{E}_i^* &= j\omega\mu\bar{H}_i^* \\ \bar{\nabla} \times \bar{H}_i^* &= j\omega\epsilon\bar{E}_i^* \end{aligned}\right\} \tag{53}$$

with

$$\bar{n} \times \bar{E}_i(r) = 0 \quad \text{on} \quad S_i - S_i' - S_i''. \tag{54}$$

We consider the following:

$$\begin{aligned} \int_V \bar{\nabla} \cdot (\bar{E}_i \times \bar{H}_{i,n}^*) d\tau &= \int_{V_i} [\bar{H}_{i,n}^* \cdot \bar{\nabla} \times \bar{E}_i - \bar{E}_i \cdot \bar{\nabla} \times \bar{H}_{i,n}^*] d\tau \\ &= \int_{V_i} [-j\omega\mu\bar{H}_i \cdot \bar{H}_{i,n}^* + j\omega_{i,n}\epsilon\bar{E} \cdot \bar{E}_{i,n}^*] d\tau. \end{aligned} \tag{55}$$

By using (51) and (53) and employing Gauss's theorem we have

$$\begin{aligned} \omega_{i,n} \int_{V_i} \epsilon(\bar{E}_i \cdot \bar{E}_{i,n}^*) d\tau - \omega \int_{V_i} \mu(\bar{H}_{i,n}^* \cdot \bar{H}_i) d\tau \\ = \frac{1}{j} \int_{S_i} (\bar{E} \times \bar{H}_{i,n}^*) \cdot n dS; \end{aligned} \tag{56}$$

and, similarly, by considering

$$\int_{Vi} \bar{\nabla} \cdot (\bar{E}_{i,n}^* \times \bar{H}_i) d\tau$$

and (52) we can show that

$$\omega \int_{V_i} \epsilon(\bar{E}_{i,n}^* \cdot \bar{E}) d\tau - \omega_{i,n} \int \mu(\bar{H}_{i,n}^* \cdot \bar{H}_i) d\tau = 0. \tag{57}$$

The normal mode expansions are

$$\bar{E}_i = \sum e_{i,n}\bar{E}_{i,n} \tag{58}$$

$$\bar{H}_i = \sum h_{i,n}\bar{H}_{i,n} + \bar{\nabla}\psi_i \tag{59}$$

where

$$e_{i,n} = \frac{\int_{V_i} \epsilon_n E_i \cdot \bar{E}_{i,n}^* d\tau}{\int_{V_i} \mu\bar{E}_{i,n} \cdot \bar{E}_{,in} d\tau} \tag{60}$$

and

$$h_{i,n} = \frac{\int_{V_i} \mu\bar{H}_i \cdot \bar{H}_{i,n} d\tau}{\int_{V_i} \mu\bar{H}_{i,n} \cdot \bar{H}_{i,n} d\tau}. \tag{61}$$

By using (56) and (57) we obtain

$$e_{i,n} = \frac{-j\omega_{i,n}}{\omega^2 - \omega_{i,n}^2} \frac{\int_{S_i'+S_i''} (\bar{E}_i \times \bar{H}_{i,n}^*) \cdot \bar{n}_i dS}{2W_n^{(i)}} \tag{62}$$

$$h_{i,n} = \frac{j\omega}{\omega^2 - \omega_{i,n}^2} \frac{\int_{S_i'+S_i''} (\bar{E}_i \times \bar{H}_{i,n}^*) \cdot \bar{n}_i dS}{2W_{i,n}} \tag{63}$$

since the surface integrals only have values at the slot surfaces.

Appendix B

Variational Form of the Solution

We shall start with the basic equations, (13) and (26), for the excitation of the slots. These equations combine to give

$$\begin{aligned} \frac{\partial^2\phi_i}{\partial x^2} + k^2\phi_i \\ = \frac{kZ_0 \sum_n \frac{\omega(1 \pm \cos\beta L)}{\omega^2 - \omega_{i,n}^2} j_{i,n,y'} \int_{-l/2}^{l/2} \bar{j}_{i,n,y'}^* \phi_i dx}{2W_{i,n}} \end{aligned} \tag{64}$$

for the case of a chain of identical cavities. Rewriting, we obtain

$$\begin{aligned} \frac{\partial^2\phi_i}{\partial x^2} + \frac{\omega^2}{c^2}\phi_i \\ = \frac{Z_0}{c} \sum_n \frac{(1 \pm \cos\beta L)\omega^2}{(\omega^2 - \omega_{i,n}^2)} \frac{j_{i,n,y'}}{2W_{i,n}} \int_{-l/2}^{l/2} \bar{j}_{i,n,y}^* \phi_i dx. \end{aligned} \tag{65}$$

Multiplying (65) by ϕ_i^*, and integrating once along the slot, we obtain

$$\int_{-l/2}^{l/2}\left(\frac{\partial^2\phi_i}{\partial x^2}+\frac{\omega^2}{c^2}\phi_i\right)\phi_i^*\,dx=\frac{Z_0}{c}$$

$$\cdot\sum_n\frac{\omega^2(1\pm\cos\beta L)}{2(\omega^2-\omega_{i,n}{}^2)W_{i,n}}\int_{-l/2}^{l/2}\bar{j}_{i,n,y}\phi_i^*dx\int_{-l/2}^{l/2}\bar{j}_{i,n,y}^*\phi_i dx. \quad (66)$$

Integrating by parts the second term on the left-hand side of (66) we have

$$\frac{\omega^2}{c^2}\int_{-l/2}^{l/2}\phi_i\phi_i^*dx-\int_{-l/2}^{l/2}\frac{\partial\phi_i}{\partial x}\frac{\partial\phi_i^*}{\partial x}=\frac{Z_0}{c}$$

$$\cdot\sum_n\frac{\omega^2(1\pm\cos\beta L)}{(\omega^2-\omega_{i,n}{}^2)2W_{i,n}}\int_{-l/2}^{l/2}\bar{j}_{i,n,y}\phi_i^*dx\int_{-l/2}^{l/2}\bar{j}_{i,n,y}^*\phi_i dx \quad (67)$$

since

$$\phi(\pm l/2)=0$$

Let us assume that $\phi_{0,i}$ and ω_0 solve (67). A trial function for the voltage is expressed as follows:

$$\phi_i(x)=\phi_{0,i}(x)+\epsilon_i(x). \quad (68)$$

Since the values of the voltage at the ends of the slot are known,

$$\epsilon_i(\pm l/2)=0. \quad (69)$$

Suppose, then, (67) yields a value of frequency ω for the trial function where

$$\omega=\omega_0+\Delta. \quad (70)$$

Then, it can be shown that Δ is of the second order in ϵ_i. Thus (67) is a variational form of the solution.

ACKNOWLEDGMENT

The authors are indebted to Prof. M. Chodorow for many helpful discussions.

Microphony in Waveguide*

I. GOLDSTEIN† AND S. SOORSOORIAN†

Summary—**This paper describes the mechanism of phase modulation by waveguide in the presence of a high intensity acoustic field. *X*-band rectangular was studied to determine the following:**

a) Resonant frequency in a transverse vibrational mode.

b) Means of minimizing phase modulation.

A CW radar can be represented as a microwave bridge in which the transmitted signal is compared in frequency with the received signal so that Doppler information may be obtained. Any disturbance of the bridge at the Doppler frequency will cause degradation in system sensitivity. It is our purpose to show that waveguide under a high acoustical field can definitely contribute to microphonics via the mechanism of phase modulation. This can be accomplished in many ways to a waveguide but we are primarily interested in transverse motion. The different transverse modes for the top and side waveguide walls are shown in Fig. 1.

Phase shift is accomplished by the motion of the side walls. The incremental phase shift is expressed as the following relationship for a rectangular waveguide operating in the TE_{10} mode:

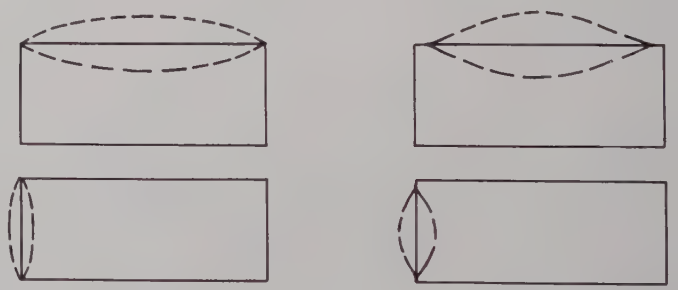


Fig. 1—First-mode vibration shapes.

$$d\phi=\frac{\pi l\lambda_g da}{2a^3},$$

$d\phi$ = incremental change of phase,
da = incremental change of side wall,
a = wide dimension of the waveguide,
l = length of the waveguide,
λ_g = guide wavelength.

Fig. 1 indicates an ideal situation of no coupling between waveguide walls. However, in actual practice there is coupling between the motion of the wide walls

* Manuscript received by the PGMTT, November 19, 1959; revised manuscript received, January 21, 1960.

† Raytheon Company, Bedford, Mass.

and narrow sides of the waveguide, as will be shown in the section on experimental results.

For the sake of analysis, the sides of the waveguide can be represented as simple beams. The resonant frequencies for the bending vibration of simple beams can be expressed in one simple form,

$$f = \frac{bc_1 t}{d^2}, \qquad (1)^1$$

where

f = frequency in cycles per second,
b = a constant dependent on shape and mode,
c_1 = the longitudinal wave velocity for the material,
t = the thickness,
d = a characteristic length.

A definition of the pin-pin and clamp-clamp boundary condition is in order. As shown in Fig. 1, the pin-pin condition is one in which there is a definite slope of the beam deflection at the boundaries. The clamp-clamp condition is one in which there is zero slope of the beam deflection at the boundaries.

In (1) for the resonant frequencies,

$b = 0.45$ for pin-pin condition,
$b = 1.01$ for the clamp-clamp condition.

Fig. 2 shows the typical acoustical environment for a large missile and the simulated conditions. Various samples of waveguide were made part of a microwave bridge and excited by a tweeter placed one inch from the broad face. Fig. 3 shows a schematic of the complete system. Power from a frequency stabilized microwave source was fed into the microwave circuit. One arm acts as the local oscillator when modulated by a 30-mc oscillator. A sideband is extracted through a tunable high Q cavity and is observed on a microwave spectrum analyzer. The other arm containing the RF carrier enters a microwave bridge where the carrier is suppressed by an adjustable phase shifter. Acoustical energy exciting the test piece at resonant frequencies will cause the bridge to unbalance because of the dimensional changes in the test piece; the side bands resulting from this phase modulation mix with the local oscillator to provide a phase-modulated 30-mc output for the IF amplifiers. It is then balance-bridge detected and fed into a video amplifier through a band-pass filter, and a visual display shows the relative amplitudes and resonant frequencies.

To calibrate the station, a ferrite modulator is substituted for the piece under test and driven at various side frequencies and levels to obtain a calibration curve. The following relationship yields the power in the sidebands:

$$P_{sB} = \frac{(\theta_{\max})^2}{16}, \qquad ^2$$

where θ = phase shift in the test arm.

[1] P. M. Morse, "Vibration and Sound," McGraw-Hill Book Co., Inc., New York, N. Y., 2nd ed.; 1948.
[2] Derived in the Appendix.

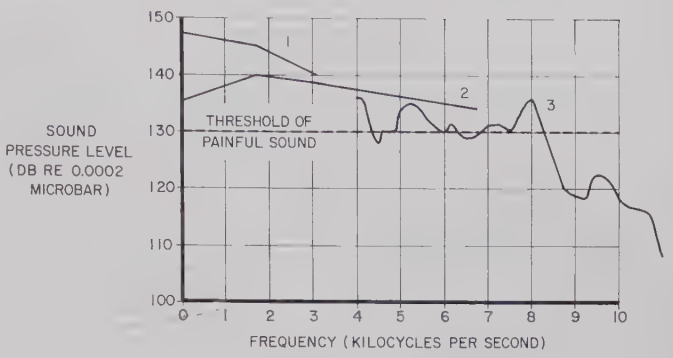


Fig. 2—Near-field spectrum of rocket noise. Curve 1, solid propellant with re-ignition of exhaust gases downstream of exit. Curve 2, solid propellant that does not re-ignite. Curve 3, sound level of test station.

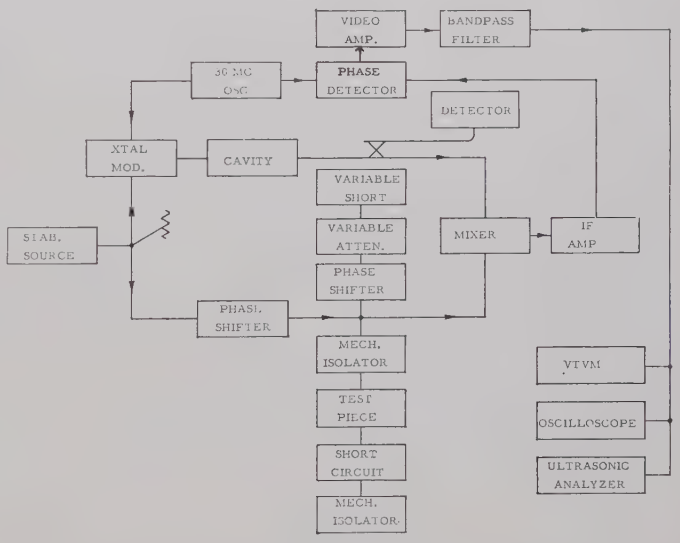


Fig. 3—Microphonic test station.

Incidentally, the sensitivity of the system is greater than 0.003 degree of phase shift.

Two areas were studied,

a) various waveguide geometries,
b) effectiveness of various coating materials on the outside of the waveguide to act as acoustical and vibrational dumping agents.

Fig. 4 shows the geometries of an X-band waveguide tested along with a simulation of the pin-pin and clamp-clamp conditions. Incidentally, the resonant frequencies of the narrow sides (0.4×0.9 ID), which are 28.4 kc pinned and 64.6 kc clamped, were not observed. Essentially then, phase modulation was caused by the motion of the wide walls (top and bottom) at their resonant frequencies coupling to the side walls. Fig. 5 shows a table of results. Note that for waveguide thicknesses greater than 0.2 inch no output was observed. Fig. 6 shows results which indicate that to a first order at least, the resonant frequencies of different waveguides were independent of length. The conclusion drawn is that truly transverse resonant effects have been observed.

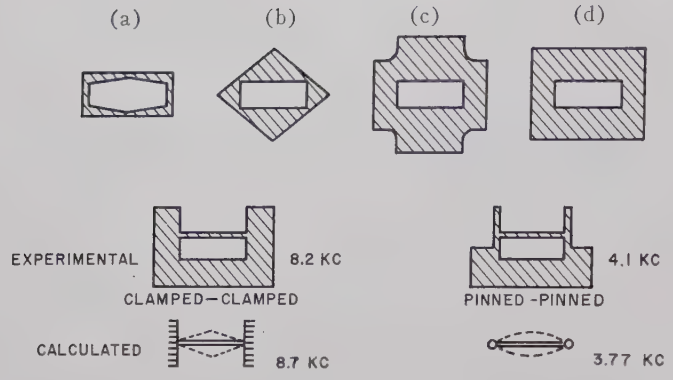


Fig. 4—Waveguide geometries.

Configuration	Wall Thickness	Measured Resonant Frequency kc	Relative Amplitude, Volts
Tapered (a)	—	6.0	0.5
Diamond (b)	—	3.5	0.5
Cross (c)	—	—	—
Heavy Wall (d)	0.300	—	—
Heavy Wall	0.200	—	—
Heavy Wall	0.150	11.7	0.015
Heavy Wall	0.125	10.2	0.075
Heavy Wall	0.100	8.4	0.40
Heavy Wall	0.075	7.3	1.50
Conventional	0.050	5.9	20+*
Light Wall	0.025	3.2	20+*

* Saturation noise level≅0.015 volt, $f=bc_1t/d^2$.

Fig. 5—Resonant frequencies and relative amplitude for various waveguide configurations. Material: brass; length 12 inches; 0.400×0.900 ID.

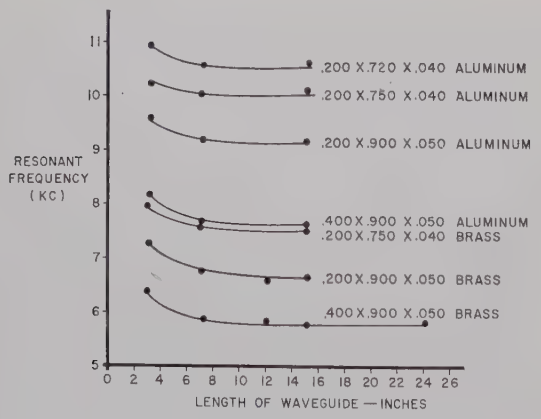


Fig. 6—Waveguide natural resonant frequencies.

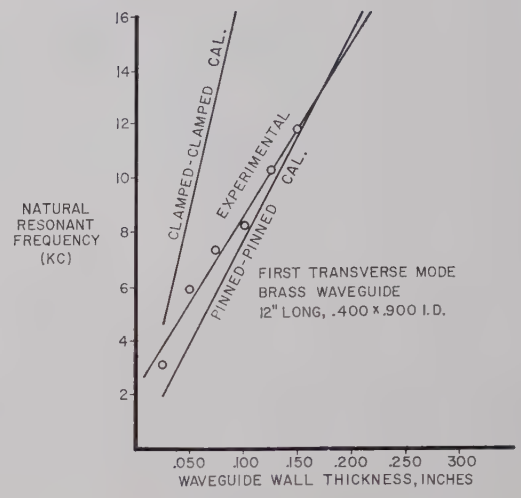


Fig. 7.

Thickness	Frequency Pin	Frequency Clamp	$\frac{f_{pin}+f_{clamp}}{2}$	$f_{Exp.}$
Inch	KC Calculated	KC Calculated	KC Calculated	
0.150	11.20	25.70	18.40	11.7
0.125	9.41	21.20	15.30	10.2
0.100	7.55	17.40	12.47	8.4
0.075	5.70	13.10	9.40	7.3
0.050	3.77	8.70	6.23	5.9
0.025	1.87	4.16	2.98	3.2

Fig. 8.

Description of Coating Materials	Resonant Frequency KC	Relative Amplitude, Volts	−DB from Uncoated Guide
Aquaplas F50/180A	7.6	0.13	43.6
Aquaplas F102A	6.4	0.20	40.0
Johns-Manville No. 678	5.3	0.33	35.6
Minnesota Mining & Mfg. EC-822	5.7	3.50	34.0
Johns-Manville No. 806	5.5	0.60	30.4
Aquaplas F100A	6.9	0.60	30.4
Minnesota Mining & Mfg. EC-244	5.6	0.90	27.0
Minnesota Mining & Mfg. EC-549	5.6	1.00	26.08
Tesamoll $\frac{1}{4}''$	5.8	3.50	15.2
Johns-Manville No. V20	5.7	5.0	12.0
Tesamoll $\frac{1}{8}''$	5.8	8.00	8.0
Stabond C-890	5.7	8.80	7.2
Johns-Manville D.B. C.T.	5.8	15.0	2.6
Noise Level≅0.015 Volt			

Fig. 9—Damping efficiencies of various coating materials of thickness $\frac{1}{8}$ inch. Brass waveguide 0.400×0.900 ID; 0.050 wall thickness; length 12 inches.

Fig. 7 shows the experimental and calculated results. A brief discussion is in order concerning the calculated vs the experimental values of resonant frequencies.

As shown in Fig. 8, for thinner wall waveguides the resonant frequency is approximately the average of pin-pin and clamp-clamp frequencies. With increasing thicknesses, the calculated pin-pin frequencies appear to yield a better agreement with experiment.

Fig. 9 shows a table of results using different coatings on the outside of the waveguide. Unfortunately the coating which appeared to be the best suffered from the following characteristics:

a) variation in acoustic attenuation with temperature,
b) solubility in water.

This area is a promising one and more research is needed in acoustic damping materials. This bridge technique has been used as a general tool for detecting resonances in microwave components and particularly high power microwave tubes. Mention should be made of a technique that can be used in most laboratories, that of an air blast. An air hose using a $\frac{1}{4}$-inch nozzle and 40 pounds of air pressure has been used and found sufficient to excite resonances up to 50 kc.

Conclusion

It has been shown that a high acoustic field can produce phase modulation of a waveguide, which can seriously degrade any system where Doppler information is desired. Methods of minimizing this effect have been demonstrated.

There is still a need for getting the best microwave structures for components such as magic "T's," hybrids, mixers, directional couplers, etc., which have minimum phase variation in an acoustical field.

Appendix

If we let the input voltage to the bridge be $V_{in} = E_n \sin \omega t$ in the null arm, we have

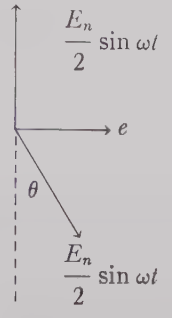
$$e = \frac{E_n}{2} \sin \omega t \sin \theta,$$

where

$$\theta = \text{phase shift in test arm} = k \sin \omega_p t,$$

$$e = \frac{E_n}{2} \sin \omega t \sin (k \sin \omega_p t)$$

$$= \frac{E_n}{2} \sin \omega t \left[2 \sum_{n=0} J_{(2n+1)}(k) \sin (2x + 1)\omega_p t \right].$$



With expansion and modification of this last equation, the following results:

$$\frac{\text{First sideband power}}{\text{Incident power}} = P_{sB} = \frac{[J_1(k)]^2}{4}.$$

For small values of k, $J_1(k) \approx K/2 \approx \theta$,

$$\therefore \; P_{sB} = \frac{(\theta_{max})^2}{16}.$$

Correspondence

Discussion of Line Width and Gyromagnetic Ratio*

A considerable amount of work has been reported on the measurement and interpretation of line width and gyromagnetic ratio of microwave ferrite materials. Most frequently, measurements are made on a small sphere placed in a resonant cavity.[1-4] It is the purpose of this letter to summarize some of the significant results of this work, particularly from the point of view of an engineer using ferrites.

Ferromagnetic Resonance

Ferrites at microwave frequencies have losses that depend on the material, its geometry, the frequency, and the magnetic biasing field. Consider first that the ferrite is unbounded. For a fixed frequency, the curve of absorption of energy by the ferrite vs biasing field exhibits a resonance, and is similar to the curve of current vs capacitance of a series LCR circuit. Resonance occurs at a field H given by

$$H = \omega/\gamma \tag{1}$$

where ω is the angular test frequency and γ is the gyromagnetic ratio. This is a ferromagnetic resonance at which the susceptibility (ratio of RF magnetization to RF magnetic field) has a resonant rise.

Line width is a measure of the sharpness of the resonant rise of absorption: the smaller the line width, the sharper the resonant rise. In this way, line width is analogous to the dissipation factor of an LCR circuit. Line width is defined as the separation of the two magnetic-field values at which the loss component of susceptibility is one half its maximum value.

Nature of Absorption Resonance in the Case of a Sphere

In practice, measurements of line width and gyromagnetic ratio are made using a small sphere in a resonant cavity. The maximum energy absorbed by the cavity, and hence the greatest decrease in cavity Q, will occur at an externally applied magnetic field determined from (1). However, this is not a case of ferromagnetic resonance, since that is determined by the internal biasing field, and the internal biasing field in the sphere is smaller than the applied field because of demagnetizing effects. Hogan shows that the resonant rise in absorption is due to a resonant buildup of the RF fields at the bias determined by (1).[5] However, the susceptibility is relatively small, much smaller than the susceptibility at ferromagnetic resonance. Susceptibility in this letter will always be taken to mean the internal or intrinsic susceptibility; *i.e.*, the ratio of magnetization to internal RF fields.

It is instructive to consider the formula for the change in the resonant frequency of the cavity as a function of the intrinsic permeability, μ_i, of the sample material. (Circular polarization is assumed.) $\mu_i = 1+\mu+k = 1+\mu'+k'-j(\mu''+k'')$ where μ and k are the diagonal and off-diagonal terms in the susceptibility tensor.[2,6]

$$\frac{\Delta F}{F} = A\,\frac{\mu_i - 1}{\mu_i + 2}\,. \tag{2}$$

The frequency F is complex, and A is a geometric constant. This formula shows that the greatest change in complex frequency occurs when μ_i approaches -2. Also, note that ΔF is relatively small for very large μ_i.

From the above discussion, it might appear that the line width as measured in a sphere is only indirectly related to the intrinsic line width that would be measured in an infinite medium. However, if we assume the simple mechanism of loss implied in the Landau-Lifschitz formulation and disregard the effects of porosity and anisotropy, the line width as measured in the sphere will be identical to the line width that would be measured in an infinite medium.[7]

Nature of Absorption in the Case of a Rod, Disk

As in the case of the sphere, the resonance in the absorption of a long thin rod will be due to a resonant buildup of the RF fields in the sample rather than to a ferromagnetic resonance. Also, under the assumptions listed above, the line width in this case will be the same as in the infinite medium.[7]

In the case of a thin disk, the resonance of the absorption will be a true ferromagnetic resonance. The internal biasing field at resonance will be the same as in the case of the infinite medium, and the line width will be the same as in the infinite medium even if porosity and anisotropy are taken into account. However, despite this advantage for thin disks, measurements are usually made on spheres principally because of the fact that the spheres are the only true ellipsoids that can readily be prepared.

Effects of Porosity and Anisotropy

Porosity and anisotropy cause local variations in the magnetic bias within the material, and this causes a broadening of the line width. They also cause the measured line width to have a different shape from the simple symmetrical shape predicted by the L-L formulation, since the L-L formulation disregards porosity and anisotropy. Since, for resonance at a given frequency, the magnitude of the internal bias field in the case of the sphere, rod, and disk are different one from the other, and since the effects of anisotropy and porosity depend to some extent on the magnitude of the bias field, the measured line width will depend on the sample shape.[8]

It is important to note that magnetic losses far off resonance are often much lower than the loss which would be predicted on the basis of the measured line width. This has been observed experimentally by Rowen and Von Aulock.[2] Far off resonance, the losses are approximately those that would have been predicted on the basis of a line width that would exist if there were no broadening caused by anisotropy and porosity. The reason for this may be understood by the radio engineer from the following circuit analogy. Consider two tuned amplifiers. In the one, the tuned circuits in each of the stages are identical. In the second amplifier, the tuned circuits are similar to those of the first, but each circuit is tuned to a slightly different frequency. The second amplifier will have a much broader resonance than the first. However, far off resonance, the response of both amplifiers will be close.

Surface Roughness, Spin Waves

For materials with very narrow intrinsic line widths, such as single crystals of yttrium iron garnet, the surface roughness of the sphere being tested can have a predominant role in determining line width. Also, for such materials, losses caused by spin wave coupling become important.[9] Since the losses due to the spin wave coupling depend in part on the demagnetizing factors, the shape of the specimen may appreciably affect the line width.

Sample Size of the Sphere

When we use perturbation techniques, the sample must be sufficiently small that the external RF field it sees is essentially

* Received by the PGMTT, December 3, 1959.

[1] J. O. Artman and P. E. Tannenwald, "Measurement of susceptibility tensor in ferrites," *J. Appl. Phys.*, vol. 26, pp. 1124–1132; September, 1955.

[2] J. H. Rowen and W. Von Aulock, "Measurement of the dielectric and magnetic properties of ferromagnetic materials at microwave frequencies," *Bell Sys. Tech. J.*, vol. 36, pp. 427–448; March, 1957.

[3] E. G. Spencer, R. C. LeCraw, and F. Reggia, "Measurement of the microwave dielectric constants and tensor permeabilities of ferrite spheres," Proc. IRE, vol. 44, pp. 790–800; June, 1956.

[4] R. A. Waldron, "Resonant cavity methods of measuring ferrite properties," *Brit. J. Appl. Phys.*, vol. 9, pp. 439–442; November, 1958.

[5] C. L. Hogan, "Elements of nonreciprocal microwave devices," Proc. IRE, vol. 44, pp. 1345–1368; October, 1956. See p. 1353.

[6] R. A. Waldron, "Ferrites in resonant cavities," *Brit. J. Appl. Phys.*, vol. 7, p. 114; March, 1956.

[7] A. D. Berk, "Dependence of the ferromagnetic resonance line width on the shape of the specimen," *J. Appl. Phys.*, vol. 28, pp. 190–192; February, 1957.

[8] E. G. Spencer, L. A. Ault, and R. C. LeCraw, "Intrinsic tensor permeabilities of ferrite rods, spheres, disks," Proc. IRE, vol. 44, pp. 1311–1317; October, 1956.

[9] C. R. Buffler, "Ferrimagnetic resonance near upper limit of the spin wave manifold," *J. Appl. Phys.*, suppl. to vol. 30, pp. 172S–175S; April, 1959.

uniform. Also, the sample must be sufficiently small that there are no "retardation" or propagation effects within the sample. The propagation constant within the sample can be calculated from its dielectric constant and taking its permeability as -2. [See (2).] Tompkins and Spencer have derived a formula to take sample size into account.[10] Their formula when expanded in a Taylor series, becomes

$$\frac{\Delta F}{F} = A \frac{(\mu_i - 1)\left(1 - \frac{X^2}{10}\right)}{(\mu_i + 2) - X^2\left(\frac{\mu_i - 1}{10} + \frac{1}{2}\right)} \cdot \quad (3)$$

$X = 2\pi r/\lambda_0\sqrt{2\epsilon'}$ where r is the radius of the sample, ϵ' is the real part of the dielectric constant of the ferrite (approximately 10), amd λ_0 is the free space wavelength. According to this equation, the error in gyromagnetic ratio measured at 3000 megacycles for a sample diameter of 0.240 inch, or at 9000 megacycles for a sample diameter of 0.080 inch, will be about two per cent. The error in line width is comparable.

Spencer *et al.* give experimental data to show that the line width of $R-1$ measured at X band is independent of sample size for diameters ranging from 25 to 60 mils.[11] Stinson shows that the line width of polycrystalline YIG measured at X band is independent of sample size for diameters ranging from 40 to 90 mils.[12] Stinson attributes the independence of sample diameter to his use of a cross-guide coupler instead of a resonant cavity, and refers to an article by Artman to show that sample size has a strong effect on measurement of line width when a cavity is used.[13] However, as noted above, Tompkins and Spencer have derived an equation for measurement in a cavity which shows only a very small dependence on diameter for the range of diameters covered by Stinson. Tompkins and Spencer discuss the discrepancy between their equation and that of Artman, and this writer believes that Tompkins and Spencer are correct.

The writer appreciates helpful comments received from Dr. R. C. LeCraw of Bell Telephone.

I. Bady
U. S. Army Signal Res. and Dev. Lab.
Fort Monmouth, N. J.

[10] J. E. Tompkins and E. G. Spencer, "Retardation effects caused by ferrite sample size on the frequency shift of a resonant cavity," *J. Appl. Phys.*, vol. 28, pp. 969–974; September, 1957.
[11] E. G. Spencer, R. C. LeCraw, and L. A. Ault, "Note on cavity perturbation theory," *J. Appl. Phys.*, vol. 28, pp. 130–132; January, 1957.
[12] D. C. Stinson, "Experimental techniques in measuring ferrite line widths with a cross-guide coupler," 1958 WESCON Convention Record, pt. 1, pp. 147–150.
[13] J. O. Artman, "Effects on the microwave properties of ferrite rods, discs, and spheres," *J. Appl. Phys.*, vol. 28, pp. 92–98; January, 1957.

Plotting Impedances with Negative Resistive Components*

The plotting of impedances with negative resistive components on some sort of inverted Smith chart is becoming more common.[1] This note suggests standardizing on a particular form, for psychological reasons. The suggested form is represented by "Γ" $= -1/\Gamma$, where Γ is the actual complex reflection coefficient, and "Γ" is the value plotted on the chart. The corresponding impedance relation is "Z/Z_0" $= -Z_0/Z$. The advantages claimed for this particular form are:

1) The transformation is analytic as opposed to the one mentioned by Stock and Kaplan.[1]

2) If both negative and positive resistances are being plotted on two Smith charts, the result, as shown in Fig. 1, looks like the representation of the world on the covers of some atlases. It fits well with the concept of projection on the unit sphere.

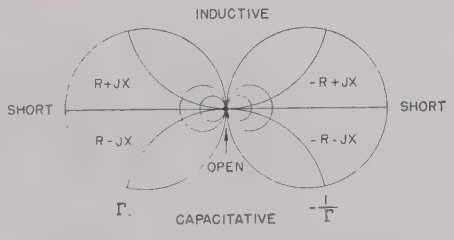


Fig. 1—Double SMITH-HTIMS chart.

3) It follows from 1) that if impedance is being plotted as a function of real frequency, then points of stability or instability as indicated by complex frequency will fall on the same side of the curve on both charts.

A possible disadvantage is the opposite sense of rotation of the two charts for transmission-line calculations, but this seems natural and easy to remember for two circles in "contact."

R. L. Kyhl
Electrical Eng. Dept. and
Res. Lab. of Electronics
Mass. Inst. Tech.
Cambridge, Mass.

Comments on "The Design of Ridged Waveguide"*

An article by Hopfer,[1] which appeared in 1955, takes into account the step discontinuity susceptance in the computation of the cutoff frequencies of ridged-waveguide. Cutoff frequencies are computed utilizing the transverse resonance method. Values of the normalized step susceptance that were used in computing the cutoff frequencies were taken from published data in the Waveguide Handbook.[2]

It seems that this procedure for determining the step susceptance is questionable. The transverse resonance method as applicable to ridged waveguide entails computing the circuit parameters of parallel plane transmission lines. Consequently, the step discontinuity susceptance should be computed as a step in a parallel plane transmission line[3,4] rather than as a step in rectangular waveguide.

Richard M. Kurzrok
RCA Surface Communications
Systems Laboratory
75 Varick Street
New York 13, N. Y.

* Received by the PGMTT, January 18, 1960.
[1] D. J. R. Stock and L. J. Kaplan, "The representation of impedances with negative real parts in the projective chart," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, p. 475, October, 1959; L. J. Kaplan and D. J. R. Stock, "An extension of the reflection coefficient chart to include active networks," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, pp. 298–299; April, 1959. R. M. Steere, "Novel applications of the Smith chart," *Microwave J.*, vol. 3, pp. 97–100; March, 1960.

* Received by the PGMTT, January 18, 1960.
[1] S. Hopfer, "The design of ridged waveguide," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 20–29; October, 1955.
[2] N. Marcuvitz, "Waveguide Handbook," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 10, pp. 399–402; 1951.
[3] S. B. Cohn, "Properties of ridge waveguide," Proc. IRE, vol. 35, pp. 783–788; August, 1947.
[4] J. R. Whinnery and H. W. Jamieson, "Equivalent circuits for discontinuities in transmission lines," Proc. IRE, vol. 32, pp. 98–116; February, 1944.

Contributors

David K. Adams was born in Portland, Ore., on August 24, 1930. He received the B.A. degree in physics in 1952 from Reed College in Portland, and the M.A. degree in physics in 1953 from the University of British Columbia, Vancouver, B. C., Can.

D. K. Adams

Mr. Adams served in the U. S. Army from 1954 to 1956 as an instructor in the Armed Forces Special Weapons Project, at Sandia Base, N. Mex., and worked for the Sandia Corporation until 1957. He then returned to graduate studies at the University of Michigan, Ann Arbor, and is currently a doctoral candidate in electrical engineering; he is also a research associate with the Electronic Defense Group of the University, working on solid-state devices.

Mr. Adams is a member of Sigma Xi.

❖

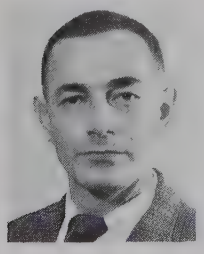
Matthew A. Allen (S'56–M'59) was born in Edinburgh, Scotland, on April 27, 1930. He attended Edinburgh University from 1947 to 1952 receiving the B.Sc. (Hons.) degree in physics in 1951. He served two years with the REME Corps of the British Army as a radar officer. He attended Stanford University, Stanford, Calif, from 1955 to 1959, receiving the Ph.D. degree in physics in 1959. Since 1955 he has been at the Microwave Laboratory at Stanford University and is presently engaged in research on high-power microwave tubes and plasma physics.

M. A. Allen

Dr. Allen is a member of the American Physical Society and Sigma Xi.

❖

Helmut M. Altschuler (S'47–A'49–M'54–SM'55) was born in Mannheim, Germany, on February 13, 1922. He received the B.E.E. and M.E.E. degrees from the Polytechnic Institute of Brooklyn, N. Y., in 1947 and 1949, respectively, and is presently continuing graduate studies there.



H. M. Altschuler

He held a research fellowship at the Polytechnic Institute of Brooklyn in 1947 and 1948, and since then he has been employed there, presently as research associate. His work has been chiefly concerned with the development of impedance meters, microwave measurement techniques, and equivalent network representations.

Mr. Altschuler is a member of Sigma Xi and Eta Kappa Nu.

❖

Robert Bawer (S'47–A'50) was born on February 12, 1925, in Ellenville, N. Y. He first entered the University of Florida, Gainesville, in 1942, and after a two-year interruption of his studies for military service, received the B.E.E. degree in 1947. From 1947 to 1949, he served as a research assistant at the Laboratory for Insulation Research, Massachusetts Institute of Technology, Cambridge. He received the M.S. degree from M.I.T. in 1949, and in that year joined the staff of Melpar, Inc., Falls Church, Va., where he was engaged as a project engineer in the design and development of multichannel communication equipment, railroad signaling and control devices, and microwave systems and components.

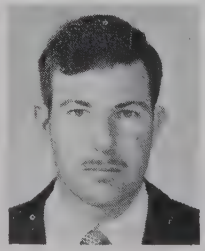
R. Bawer

In 1956, he joined Emerson Research Laboratories in Washington, D. C., as technical consultant in the fields of microwave and countermeasure systems. In 1958, he accepted his present position as Assistant Director for Research and Development at Aero Geo Astro Corp., Alexandria, Va.

Mr. Bawer is a member of Phi Kappa Phi, Sigma Tau, and Sigma Xi.

❖

Robert E. Collin (M'54) was born in Donalda, Alberta, Canada, on October 24, 1928. He received the B.S. degree in engineering physics from the University of Saskatchewan, Saskatoon, Canada, in 1951. The following two and a half years were spent in graduate work at Imperial College, London, England, from which he received the Ph.D. degree and the diploma of Imperial College in 1954.



R. E. Collin

Upon returning to Canada, he worked from 1954 to 1958 at the Canadian Armament Research and Development Establishment, Quebec. Since 1958, he has been on the professorial staff of the Electrical Engineering Department at Case Institute of Technology, Cleveland, Ohio.

Dr. Collin is a member of Sigma Xi.

Luther Davis, Jr. (SM'55) was born on July 12, 1922, in Mineola, N. Y. He received the B.S. degree in 1942 and the Ph.D. degree in physics in 1949, both from the Massachusetts Institute of Technology, Cambridge.

L. Davis, Jr.

During World War II, he was on the staff of the M.I.T. Radiation Laboratory. In 1949, he joined the Research Division of the Raytheon Company, Waltham, Mass., where he has been engaged in research in solid-state physics.

Dr. Davis is a member of the American Physical Society.

❖

Frans C. de Ronde was born in Schiedam, The Netherlands, on June 20, 1923. He received the degree in electrical engineering in 1953 from the Technische Hogeschool of Delft, The Netherlands.

F. C. de Ronde

He joined the Philips Research Laboratories at Eindhoven, The Netherlands in 1952, where he is concerned with research on microwave techniques and components.

❖

William J. Getsinger (S'48–A'50–M'55) was born in Waterbury, Conn., on January 24, 1924. He received the B.S. degree from the University of Connecticut, Storrs, in 1949 and the M.S. degree from Stanford University, Stanford, Calif., in 1959, both in electrical engineering.

W. J. Getsinger

In 1949–1950 he was employed by the Crystal Research Laboratories in Hartford, Conn., working on crystal filters and measuring equipment. During 1950–1951 he was an engineer with the Technicraft Laboratories, Thomaston, Conn., working on waveguide components and measurements. In 1951 he became a development engineer with the Westinghouse Electric Company, Baltimore, Md. In 1953 he returned to the Technicraft Laboratories, Inc., as a design and project engineer, working on waveguide components and assembly design. In Sep-

tember, 1957, Mr. Getsinger joined the staff of Stanford Research Institute, Menlo Park, Calif. where he has been working on stripline components and antennas.

Mr. Getsinger is a member of Tau Beta Pi and Eta Kappa Nu.

❖

Irving Goldstein (A'48–M'55) was born in Worcester, Mass., on December 19, 1920. He received the B.S. degree in electrical engineering from Worcester Polytechnic Institute, Worcester, Mass., in 1947. He served in the Army Signal Corps during World War II. He attended Brooklyn Polytechnic Institute, Brooklyn, N.Y., in 1943–1944 as a member of ASTP studying electrical engineering. He pursued graduate studies at M.I.T., Cambridge, Mass. in 1947–1948.

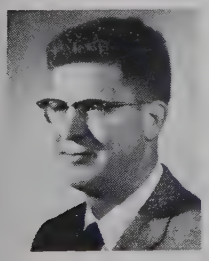
I. Goldstein

He joined Raytheon Manufacturing Company in 1947 as an electronics engineer. In 1948 he joined the microwave laboratory as a design and development engineer of components used in radar, beacons, countermeasure systems, and relay equipment. In 1954 he transferred to the Missile Systems Division as a senior engineer and then as section head of the microwave components group. At present he is manager of the Solid State Physics Branch of the Advanced Development Department at Raytheon Missile Systems Division.

Mr. Goldstein is a member of EIA.

❖

Richard C. Honey (S'48–A'53–SM'57) was born in Portland, Ore., on March 9, 1924. He received the B.S. degree from the California Institute of Technology, Pasadena, in 1945; the degree of Electrical Engineer from Stanford University, Stanford, Calif., in 1950; and the Ph.D. degree in electrical engineering from Stanford University in 1953.



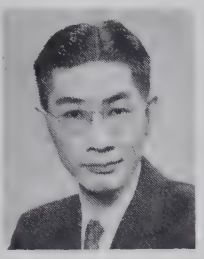
R. C. Honey

He served in the U. S. Navy as a radio technician from 1943 to 1946, and from 1948 to 1952 he was a member of the microwave oscillator project at Stanford Electronics Research Laboratory. Since 1952 he has been on the staff of Stanford Research Institute, Menlo Park, Calif., where he has worked on microwave antenna systems and components and on parametric amplifiers.

Dr. Honey is a member of the Scientific Research Society of America and Sigma Xi.

❖

Ming-Kuei Hu (S'48–A'51–SM'56) was born on May 25, 1918, in Anhwei, China. He received the B.E.E. degree from National Central University, Chungking, China, in 1941, and the Ph.D. degree from Oregon State College, Corvallis, in 1951. In the same year, he became a research assistant professor in electrical engineering at Syracuse University, Syracuse, N. Y., where he is now a senior research engineer. He has done research in high-voltage discharge phenomena, electromagnetic theory, and antenna studies.

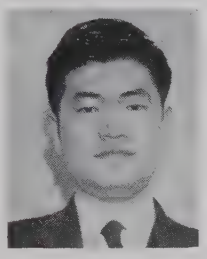


M.-K. Hu

Dr. Hu is a member of Sigma Xi, Pi Mu Epsilon, the American Physical Society, and the Association for Computing Machinery.

❖

Koryu Ishii (M'55) was born in Tokyo, Japan, on March 18, 1927. He received the B.S. degree in electrical engineering from Nihon University, Tokyo, in 1950, and the M.S. and Ph.D. degrees in electrical engineering in 1957 and 1959, respectively, both from the University of Wisconsin, Madison.



K. Ishii

From 1949 to 1956, he worked on research of microwave circuits and amplifiers and instructed students at the Nihon University.

From 1956 to 1959, he worked on research of the noise figure of the reflex klystron amplifiers and cascaded reflex klystron amplifiers at the University of Wisconsin. Since 1959, he has been engaged in establishing a millimeter-wave laboratory at Marquette University, Milwaukee, Wis., where he is an assistant professor.

Dr. Ishii is a member of Sigma Xi, ASEE, and the Institute of Electrical Communication Engineers of Japan.

❖

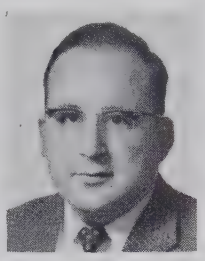
E. M. T. Jones (S'45–A'51–M'55–SM'56), for a photograph and biography, please see p. 483 of the October, 1959, issue of these Transactions.

❖

Gordon S. Kino (S'52–A'54) was born on June 15, 1928, in Melbourne, Australia. He received the B.Sc. degree in mathematics in 1948 and the M.Sc. degree in mathematics in 1950, both from London University, England.

He joined the Mullard Radio Valve Co., Salfords, Surrey, England, in 1947 as a managerial apprentice. In 1948, he became a member of the Mullard Vacuum Physics Laboratory, where he was engaged in research on microwave triodes, traveling-wave tubes, and klystrons.

From 1951 to 1955, he was employed as a research assistant at the Electronics Laboratory of Stanford University, Stanford, Calif., where he carried out research on traveling-wave tubes and electromagnetic theory, obtaining the Ph.D. degree in 1955 from Stanford University. He became a member of the technical staff of the Bell Telephone Laboratories, Murray Hill, N. J. in December, 1955, where he was associated with the Electron Tube Development Department and carried out research on magnetrons. He returned to Stanford in February, 1957, and is now a research associate in the Microwave Laboratory, where he is in charge of a plasma physics research program.



G. S. Kino

Dr. Kino is a member of Sigma Xi.

❖

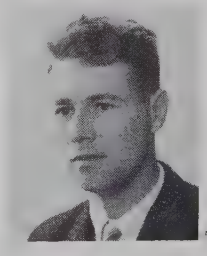
David S. Lerner (S'52–A'54–M'59) was born in Brooklyn, N. Y., on April 27, 1934. He received the B.E.E. and M.E.E. degrees from the Polytechnic Institute of Brooklyn in 1954 and 1959, respectively.

D. S. Lerner

In 1954 he joined the staff of Wheeler Laboratories, Inc., Smithtown, N. Y., where he has worked on monopulse radar antennas and artificial dielectrics.

❖

Oscar W. Memelink was born in Blinju, Indonesia, on April 11, 1927. He received the degree in physical engineering in 1952 from the Technische Hogeschool of Delft, The Netherlands.



O. W. Memelink

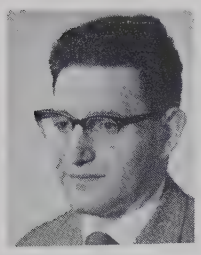
From 1952 to 1955, he was engaged in work on photoconducting compounds for the Institute of Applied Scientific Research, The Hague. Since 1955, he has been with Philips Research Laboratories at Eindhoven, The Netherlands, where he is concerned with transistor physics.

❖

H. J. G. Meyer was born in Bielefeld, Germany on August 12, 1924. In 1950, after receiving the degree in theoretical physics from the University of Amsterdam, The Netherlands, he joined Philips Research Laboratories, Eindhoven, The Neth-

erlands, where he worked on various problems in the theory of solids.

He received the Ph.D. degree from the University of Amsterdam in 1956. Presently he is engaged in a variety of investigations in the field of semiconductors as microwave resonance, infrared absorption, etc.

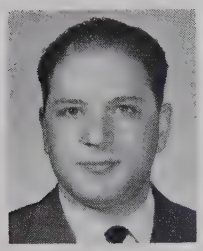


H. J. G. MEYER

Dr. Meyer is a member of the Dutch and the American Physical Society.

❖

Ugo Milano (A'57–M'60) was born in San Severo, Italy, on December 6, 1925. He received the degree in electrical engineering from the University of Bologna, Italy, in 1950 and the Diploma of Specialization in Telecommunication from the Superior Institute of Telecommunications, Rome, Italy, in 1952.



U. MILANO

He remained at the latter institution from 1952–1957, doing research in the microwave field of telecommunications. In May, 1958, he joined the Research Division of the Raytheon Company, Waltham, Mass., where he has been working on ferrite microwave devices and nonlinear phenomena in ferrites.

❖

A. A. Oliner (M'47–SM'52), for a photograph and biography, please see page 126 of the January, 1960 issue of these TRANSACTIONS.

❖

Joseph H. Saunders, for a photograph and biography, please see page 126 of the January, 1960 issue of these TRANSACTIONS.

❖

Laszlo Solymar was born in Budapest, Hungary, on January 24, 1930. He received the Diploma of Electrical Engineering from Technical University of Budapest in 1952.

L. SOLYMAR

From 1952 to 1953 he was a faculty assistant at the Technical University. From 1953 to 1956 he was employed as a research engineer at the Research Institute of Telecommunication, Budapest, where his work concerned antenna theory and design. He obtained a higher Hungarian degree in 1956. Since December, 1956, he has worked on various phases of microwave transmission at Standard Telecommunication Laboratories, Ltd., formerly in Enfield and presently in Harlow, England, where he is a research engineer.

❖

Souren A. Soorsoorian (A'50–M'52) was born in Newburyport, Mass., on October 12, 1925. He received the B.S. degree in electrical engineering from Worcester Polytechnic Institute, Worcester, Mass., in 1946 and the M.S. degree in electrical engineering from Northeastern University, Boston, Mass., in 1955.

S. A. SOORSOORIAN

From 1943 to 1947 he served in the U. S. Navy in the V-12 program and also as a Communication Officer. In 1947 he joined the Raytheon Company, Waltham, Mass., and was concerned with the design of antenna phasing equipment and microwave components. In 1951, he was recalled by the U. S. Navy and spent two years as an electronics instructor at the Naval School of Mine Warfare at Yorktown, Va.

From 1953 to 1955, he was a staff member at Lincoln Laboratory, Lexington, Mass., where he worked on instrumentation problems in the field of tropospheric scatter propagation. In 1955, he joined the Microwave Development Department of Raytheon's Missile Systems Division as a microwave engineer, and in 1956 became a section head. He is currently engaged in the development of microwave components and systems.

Mr. Soorsoorian is a member of the AIEE.

❖

Jean G. Van Bladel (M'54–SM'56) was born in Antwerp, Belgium, on July 24, 1922. He received the degree of Electrical Engineer from Brussels University, Brussels, Belgium in 1947 and the Ph.D. degree in electrical engineering from the University of Wisconsin, Madison, in 1950.

J. VAN BLADEL

From 1950 to 1954 he was head of the Radar Department, Manufacture Belge de Lampes et Matériel Electronique Factories, Brussels. From 1954 to 1956 he was associate professor of electrical engineering at Washington University, St. Louis, Mo., and since 1956 has held the same position at the University of Wisconsin. He is a consultant with the Midwestern Universities Research Association Accelerator Project.

Dr. Van Bladel is a member of the AIEE, Sigma Xi, and Eta Kappa Nu.

Harold A. Wheeler (A'27–M'28–F'35) was born in St. Paul, Minn., on May 10, 1903. He received the B.S. degree in physics from George Washington University, Washington, D. C. in 1925, and did postgraduate study in physics at The Johns Hopkins University, Baltimore, Md., until 1928.

H. A. WHEELER

He was employed by the Hazeltine Corporation from 1924 to 1946, advancing to vice-president and chief consulting engineer. In 1959, he resumed some activity with this company as a vice-president and a director. Since 1947, his principal occupation has been as President of Wheeler Laboratories, Inc., Great Neck, N. Y., now a subsidiary of Hazeltine Corporation. In this capacity, he is directing their Great Neck and Smithtown laboratories, specializing in microwave and antennas.

His specialization in frequency selective networks dates back to a college thesis on "wave filter determinants," published in 1928. Subsequent work on wideband amplifiers for television was presented in IRE papers, which were recognized by the Morris N. Liebmann Memorial Prize in 1940. "Wheeler Monographs, Vol. I" is a collection of post-war papers dealing with special topics in network theory.

Mr. Wheeler has served the IRE in such positions as Director (1934, 1940–1945), chairman of the Standards Committee, and chairman of the Long Island Section. He is a Fellow of the AIEE, and of the Radio Club of America, an Associate Member of IEE, and a member of Sigma Xi and Tau Beta Pi.

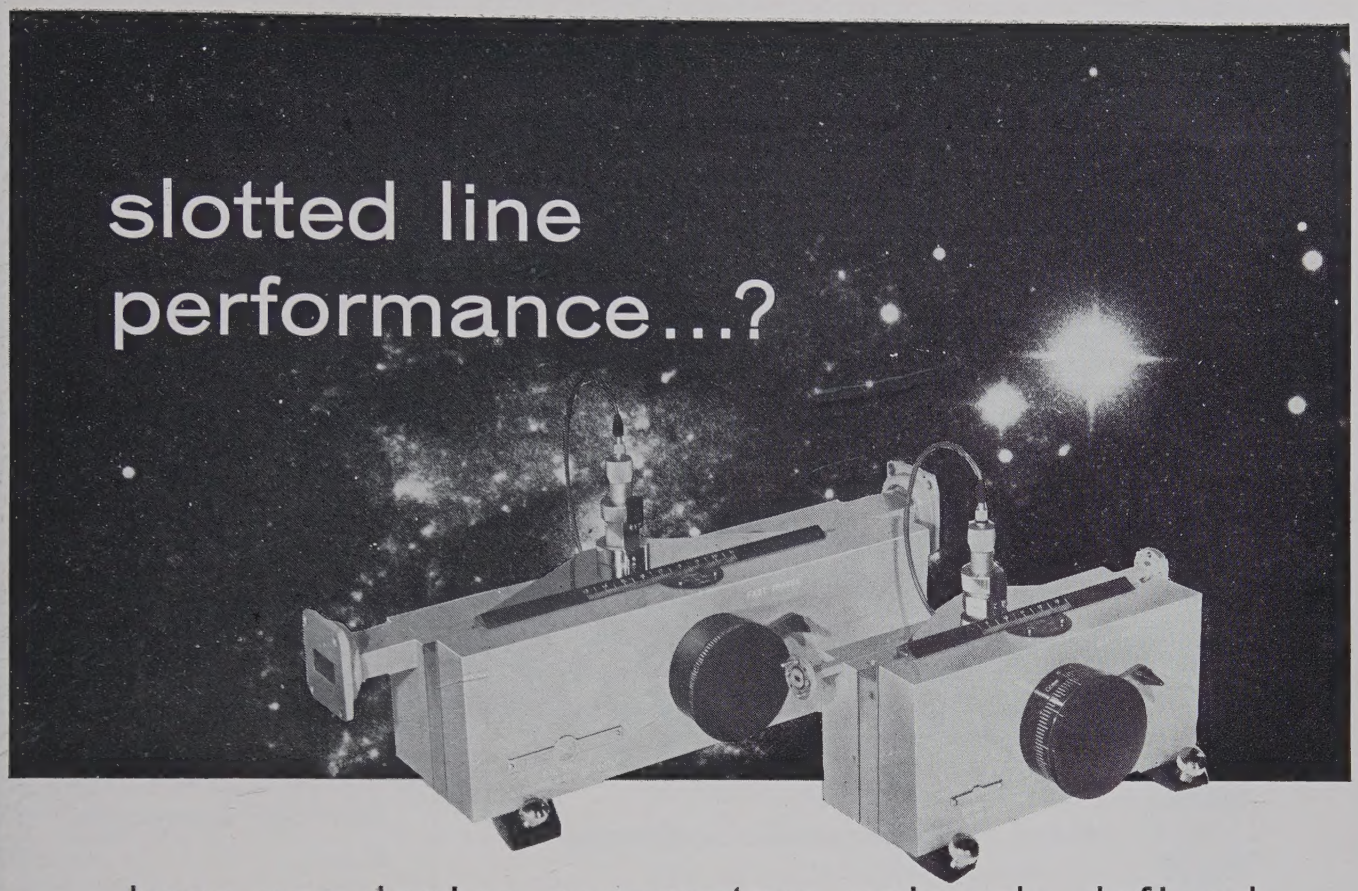
❖

John J. Wolfe (A'59) was born in McKeesport, Pa., on December 19, 1930. He served as a sonarman in the U. S. Navy from 1951 to 1953. Since 1955, he has been attending classes at the northern Virginia extension of the University of Virginia, pursuant to a degree in electrical engineering.

J. J. WOLFE

In 1956, he joined the antenna group of Melpar, Inc., Falls Church, Va., where he was employed in the research and development of microwave antennas. His fields of experience included the design of high-speed scanning antennas, polarization and phase control of wide-band antennas, and the investigation of homogeneous dielectric materials. In 1958, he joined the staff of Aero Geo Astro Corp., Alexandria, Va. Since that time, he has been engaged in the design of microwave components and investigation of broad-band antennas.

slotted line performance...?



—here are design parameters uniquely defined

Note how comprehensively D-B parameters cover all requirements of a standing wave detector.

D-B units are designed for unprecedented accuracy, repeatability, and operator convenience. They're built with high-precision parts, under close quality control. Attractively styled and fully guaranteed, they have achieved a tremendous acceptance by the industry.

Interchangeability with precision. Any D-B unit will handle adjacent frequency bands by using a different size waveguide block and probe. You can make the change in 30 seconds, with no loss whatever in alignment accuracy.

Complete range of sizes—10 models cover from 5.85 KMC to 140 KMC—or you can purchase interchangeable blocks and probes to extend the range of any model, at a saving.

Uniformity of waveguide surfaces. D-B provides a high internal surface uniformity by precision machining its millimeter waveguides, and using carefully selected precision waveguide for lower-frequency units. This construction insures a uniform path for measured waves, thus minimizing residual VSWR.

Slot excitation is negligible, resulting in minimum RF leakage—another reason why residual VSWR is very low.

Probe impedance is properly matched to the waveguide. Uniform probe penetration is provided by a fast, convenient slope adjustment, made in a few minutes.

Adequate probe travel ($\geq 1\frac{1}{2}\ \lambda_g$) available at all frequencies. Operator can read at least 3 maxima and 3 minima of VSWR.

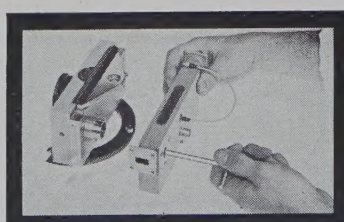
Efficient mechanical translation. D-B mechanism functions with exceptional smoothness, giving the unit a definite instrument "feel." Five point kinematic carriage suspension insures excellent linearity of probe motion.

Vernier readout on knob periphery permits reading of probe travel to .01mm without mounting of costly accessories.

Continuously variable drive ratio changes carriage travel from "vernier" to "fast"—a time saver during rapid measurements.

Direct phase readout. Phase shift may be measured accurately on the calibrated knob, which reads percentage of 180° directly.

For complete data, see your D-B Catalogue, or request folder DB-825.



Interchangeable waveguide blocks. Each realigns perfectly to probe travel in a few seconds.

DE MORNAY BONARDI®

780 SOUTH ARROYO PARKWAY • PASADENA, CALIFORNIA

NOTICE TO ADVERTISERS

•

Effective immediately the IRE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES will accept display advertising. For full details contact Tore N. Anderson, Advertising Editor, PGMTT TRANSACTIONS, 1539 Deer Path, Mountainside, N.J.

AVAILABLE BACK ISSUES OF IRE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES

	PG Members	IRE Members	Non-Members*
MTT-4, No. 3, July, 1956	1.25	1.85	3.75
MTT-5, No. 3, July, 1957	1.15	1.70	3.45
MTT-5, No. 4, Oct., 1957	1.20	1.80	3.60
MTT-6, No. 1, Jan., 1958	2.65	3.95	7.95
MTT-6, No. 2, April, 1958	2.50	3.75	7.50
MTT-6, No. 3, July, 1958	2.00	3.00	6.00
MTT-6, No. 4, Oct., 1958	2.50	3.75	7.50
MTT-7, No. 2, April, 1959	2.20	3.30	6.60
MTT-7, No. 3, July, 1959	1.65	2.45	4.95
MTT-7, No. 4, Oct., 1959	1.60	2.40	4.80
MTT-8, No. 1, Jan., 1960	2.10	3.15	6.30
MTT-8, No. 2, March, 1960	2.50	3.75	7.50

*All Libraries, Colleges and Subscription agencies may purchase at the IRE Member rate.

INSTITUTIONAL LISTINGS

The IRE Professional Group on Microwave Theory and Techniques is grateful for the assistance given by the firms listed below, and invites application for Institutional Listing from other firms interested in the Microwave field.

AIRTRON, INC., A Division of Litton Industries, 200 East Hanover Ave., Morris Plains, N.J.
Designers and Producers of Complete Line of Microwave Electronic and Aircraft Components

COLLINS RADIO CO., Texas Division, Dallas, Tex.
Complete Microwave and Transhorizon Communication Systems

ITT LABORATORIES, 500 Washington Ave., Nutley 10, N.J.
Line-of-Sight and Over-the-Horizon Microwave Systems; Test Equipment and Components

LITTON INDUSTRIES, Electron Tube Div., 960 Industrial Rd., San Carlos, Calif.
Magnetron, Klystrons, Carcinotrons, TWT's, Backward Wave Oscillators, Gas Discharge Tubes, Noise Sources

MICROWAVE DEVELOPMENT LABS., INC., 92 Broad St., Babson Park 57, Mass.
Designers, Developers and Producers of Microwave Components and Assemblies, 400 mc to 70 kmc

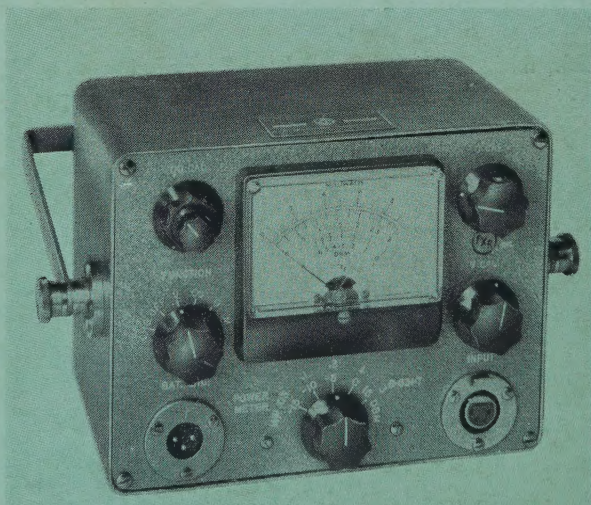
WHEELER LABORATORIES, INC., Great Neck, N.Y.; Antenna Lab., Smithtown, N.Y.
Consulting Services, Research & Development, Microwave Antennas & Waveguide Components

The charge for an Institutional Listing is $50.00 per issue or $210.00 for six consecutive issues. Applications for Institutional Listings and checks (made out to the Institute of Radio Engineers) should be sent to Tore N. Anderson, PGMTT Advertising Editor, 1539 Deer Path, Mountainside, N.J.

FXR TRANSISTORIZED
TEMPERATURE COMPENSATED
POWER METER

Measures RF and Microwave Power with Incredible Stability!

When used in conjunction with an FXR Series 218* Temperature Compensated Thermistor Head, the *Power Meter* is a hundred times more stable than comparable instruments, due to a unique method of compensating for ambient temperature changes.



Model B832T, Price $450.00

- Readings are virtually drift-free.
- Five direct reading ranges—30 microwatts to 3 milliwatts full scale.
- DC calibration at all levels.
- Range switching without recalibration.
- Self-contained re-chargeable nickel cadmium battery—use to charge ratio, 2 to 1.

* Available in coaxial and waveguide sizes from 0.01 to 40.00 KMC.

Complete details for the asking.

FXR, Inc.
Design • Development • Manufa
26-12 Borough Place • RA . 1-90
Woodside 77, N. Y. TWX: NY 43

Precision Microwave Equipment • High-Power Pulse Modulators • High-Voltage Power Supplies • Electronic Test Equipment